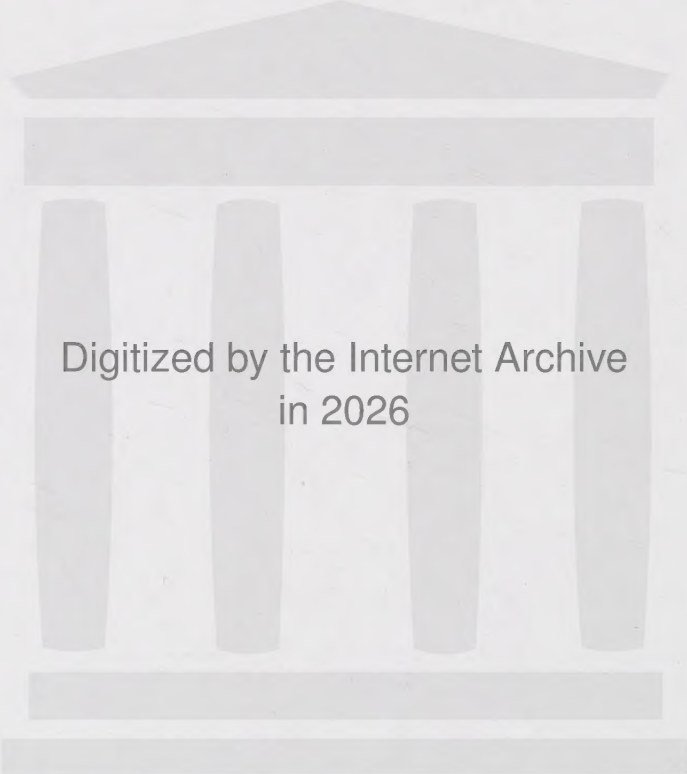


**Strategien und damit assoziierte Belastungen
von Kandidaten in mündlichen Prüfungen -
die Theorie der Gelernten Hilflosigkeit übertragen
auf das Verhaltensmuster der Prüfungsangst**

Inaugural-Dissertation
zur Erlangung des Grades
eines Doktors der Philosophie
in der
Abteilung für Philosophie, Pädagogik, Psychologie
der
Ruhr-Universität Bochum

vorgelegt von
Lothar Breitkopf

1979



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41546404_0001

**Strategien und damit assoziierte Belastungen
von Kandidaten in mündlichen Prüfungen -
die Theorie der Gelernten Hilflosigkeit übertragen
auf das Verhaltensmuster der Prüfungsangst**

Inaugural-Dissertation
zur Erlangung des Grades
eines Doktors der Philosophie
in der
Abteilung für Philosophie, Pädagogik, Psychologie
der
Ruhr-Universität Bochum

vorgelegt von
Lothar Breitkopf



Gedruckt mit Genehmigung der
Abteilung für Philosophie, Pädagogik, Psychologie
der RUHR-UNIVERSITÄT BOCHUM

Referent: Professor Dr. D. Schulte
Koreferent: Professor Dr. H. Heckhausen

Tag der mündlichen Prüfung: 5. Juli 1979

Mein besonderer Dank für die Unterstützung bis
zur Fertigstellung dieser Arbeit gilt

Herrn Professor Dr. D. Schulte,
Herrn Professor Dr. H. Heckhausen,
Frau Professor Dr. G. Krüskemper,
den Mitarbeitern der AE Klinische Psychologie,
insbesondere Herrn Dipl.-Psych. R. Künzel,
den beteiligten Prüfern und
insbesondere den Prüfungskandidaten, die
trotz ihrer eigenen Belastungen Zeit und
Arbeit für unser Projekt hatten,
sowie meiner Frau Hildegard.

Liste besonderer Abkürzungen und Kürzel

AN	Bezeichnung für den Fragebogen AN, der nach Beendigung der Prüfung, aber <u>vor</u> Bekanntgabe der Zensur von den Prüfungskandidaten bearbeitet wurde; die Bezeichnung wurde wegen des hohen Anteils an <u>Analogskalen</u> gewählt
BDI	<u>Beck-Depressions-Inventory</u> ; in dieser Arbeit: BDI-Selbststrating-Kurzform
BL	<u>Baseline</u> ; Datenerhebungszeitpunkt ca. 14 Tage vor Beginn der individuell jeweils ersten mündlichen Prüfung
BL1, BL2	Kürzel in den Tabellen für die Faktorscores der beiden Faktoren der Beschwerdeliste (BL1 = Erschöpfungssyndrom, BL2 = Angstsyndrom)
C	<u>Consequence</u> , Verstärker
EK	external-konstante Attribuierung (z. B. 'Prüfer')
EV	external-variable Attribuierung (z. B. 'Glück')
FA	<u>Faktorenanalyse</u>
FBL	<u>Freiburger Beschwerdeliste</u>
F2, F3	Bezeichnung für Fragebögen, die nach der Prüfung bearbeitet wurden (F2 = retrospektive Daten, F3 = prospektive Daten)
HD, HDn	<u>Hauptdiplom</u> , <u>Hauptdiplomand</u> , <u>Hauptdiplomanden</u>
HiS	<u>Hilflosigkeitsskala</u>
HoS	<u>Hoffnungslosigkeitsskala</u>
IK	internal-konstante Attribuierung (z. B. Fähigkeit)

IV	<u>internal-variable</u> Attribuierung (z. B. Anstrengung)
K	Kontingenz; Wahrscheinlichkeit des gemeinsamen Auftretens von Reaktion und Verstärker
KFA, KFA _n	<u>Konfigurationsfrequenzanalyse</u> , <u>Konfigurationsfrequenzanalysen</u>
K-Gruppe	Gruppe im triadischen Design, die mit kontrollierbaren Verstärkern konfrontiert wird
LH	<u>Learned Helplessness</u> , Gelernte Hilflosigkeit
N	Anzahl 'N' der Versuchspersonen
N-Gruppe	Gruppe im triadischen Design, die <u>naiv</u> in die Testphase geht
O	Organismus
P1, P2, P3, P4	fortlaufend numerierte <u>Prüfungsdatenerhebungszeitpunkte</u>
R	<u>Reaktion</u>
S ^D	<u>Diskriminativer Stimulus</u> (Reiz)
SGI	<u>Studiengewohnheitsinventar</u>
SS	<u>Strategiensammlung</u>
STAI	<u>State-Trait-Anxiety-Inventory</u>
TVD, TVD(W) TVD(H)	Testverhaltensdifferential in zwei Versionen: der prospektiven 'Werde'-Version (TVD(W)) und der retrospektiven 'Habe'-Version (TVD(H))
UK-Gruppe	Gruppe im triadischen Design, die mit <u>unkontrollierbaren</u> Verstärkern konfrontiert wird
VD, VD _n	<u>Vordiplom</u> , <u>Vordiplomand</u> , <u>Vordiplomanden</u>
VL	<u>Versuchsleiter</u>
Vp, Vp _n	<u>Versuchsperson</u> , <u>Versuchspersonen</u>

INHALTSVERZEICHNIS

	Seite
1 EINLEITUNG	1
2 THEORETISCHER TEIL	9
2.1 Prüfungen und Prüfungsbelastungen	9
2.1.1 Die Funktionen von Prüfungen aus sozialwissenschaftlicher Sicht	10
2.1.1.1 Anthropologische Konzepte	11
2.1.1.2 Soziologische, systemfunktionalistische Konzepte	15
2.1.1.3 Symbolisch-interaktionistische Ansätze	22
2.1.1.4 Ansätze der Kritischen Theorie	24
2.1.1.5 Zusammenfassung und Diskussion	30
2.1.2 Die psychologischen Komponenten der mündlichen Prüfung	37
2.1.2.1 Die Interaktion in Prüfungen als Rahmenbedingung für Prüfungsstrategien - symbolisch-kommunikative Betrachtungsebene	37
2.1.2.2 Die Symptomatik der Prüfungsbelastung - organisch-physiologische, motorisch-verhaltensmäßige und subjektive Betrachtungsebenen	40
2.1.2.3 Zusammenfassung und Diskussion	48
2.2 Die Theorie der Gelernten Hilflosigkeit - Generalisierung und Differenzierung der Theorie in drei Phasen	52
2.2.1 Phase I: Beginn mit Tierexperimenten	55
2.2.1.1 Allgemeine Darstellung der Theorie	55

	Seite
2.2.1.1.1	Die drei Aspekte der Gelernten Hilfflosigkeit: kognitives, motivationales und emotionales Defizit 60
2.2.1.1.2	Das Immunisierungssparadigma 63
2.2.1.1.3	Die Unterscheidung von drei methodischen Ansätzen 64
2.2.1.1.4	Hauptforschungsmethode: die Realisation unterschiedlicher Vorerfahrung im triadischen Design 66
2.2.1.1.4.1	Darstellung eines Beispiels 71
2.2.1.1.4.2	Transfer der Gelernten Hilfflosigkeit 74
2.2.1.2	Kritische Würdigung von Phase I 78
2.2.2	Phase II: Humanversuche - Betonung der kognitiven Komponente 84
2.2.2.1	Darstellung eines Beispiels 84
2.2.2.2	Transfer der Gelernten Hilfflosigkeit 89
2.2.2.3	Erweiterungen der Theorie 96
2.2.2.3.1	Integration des Konfliktkonzepts 96
2.2.2.3.2	Gelernte Hilfflosigkeit als Folge von Instruktionen 100
2.2.2.3.3	Die Situationskognition in Experimenten zur Gelernten Hilfflosigkeit und die Kausalattribution der Versuchspersonen 103
2.2.2.4	Zusammenfassung und Diskussion 107
2.2.3	Phase III: Humanversuche - Integration der Reaktanztheorie 111
2.2.3.1	Allgemeine Darstellung der Reaktanztheorie 111
2.2.3.2	Darstellung des integrativen Modells 113
2.2.4	Die äußere Validität der Experimente zur Gelernten Hilfflosigkeit 120

	Seite
2.3 Die Übertragung der Theorie der Gelernten Hilfllosigkeit auf die Prüfungsinteraktion	123
2.3.1 Symbolisch-kommunikative Betrachtungsebene	124
2.3.1.1 Eine kommunikationstheoretische Reanalyse der Experimente zur Gelernten Hilfllosigkeit	124
2.3.1.2 Ableitung der Hypothesen zum Verhältnis von Vorerfahrung, Prüfungsstrategie und Leistungsdaten	129
2.3.2 Organisch-physiologische, motorisch-verhaltensmäßige und subjektive Betrachtungsebene	134
2.3.2.1 Die Erklärung der Depression durch die Theorie der Gelernten Hilfllosigkeit	134
2.3.2.1.1 Passivität, kognitives Set und Zeitverlauf	135
2.3.2.1.2 Mangelnde Aggressivität	138
2.3.2.1.3 Sexuelle, soziale und Ernährungsdefizite, Störungen des Noradrenalinhaushalts und Magengeschwüre	140
2.3.2.1.4 Affekte	141
2.3.2.2 Exkurs: Zum allgemeinen Verhältnis von Angst und Depression	142
2.3.2.3 Ableitung der Hypothesen zum Verhältnis von Vorerfahrungen, Prüfungsstrategie und Prüfungsbelastungen	146
2.3.3 Integration der Betrachtungsebenen: Beziehungen zwischen Vorerfahrungen, Prüfungsstrategien, Leistungsdaten und Prüfungsbelastungen	148
2.3.4 Prüfungsstrategien: ihre Messung und Validierung	153

	Seite
3 METHODISCHER TEIL	160
3.1 Methodische Rahmenbedingungen: die Prüfung als natürliches Experiment	160
3.1.1 Selektion der Stichprobe	161
3.1.2 Einschränkende Schwierigkeiten der Forschungspraxis	165
3.1.3 Prinzipielle Planung der Datenerhebung	167
3.1.4 Zusammenfassung und Schlußfolgerung	170
3.2 Maßnahmen, die die Behauptung eines globalen Meßfehlers widerlegen sollen	172
3.2.1 Zur internen Validität der Datenerhebung: die Unterscheidung von Peripherdaten und Kerndaten	172
3.2.2 Die Instrumente zur Erhebung der Peripherdaten und die peripheren Hypothesen	174
3.2.2.1 Die Freiburger Beschwerde- liste (FBL)	174
3.2.2.2 Das Studiengewohnheitsin- ventar (SGI) zur Analyse von Arbeitsstörungen	176
3.2.2.3 Analogskalen zur Messung von Angstverläufen	178
3.2.3 Zusammenfassung	180
3.3 Meßinstrumente zur Hypothesenprüfung	181
3.3.1 Die Messung der emotionalen Belastung	181
3.3.1.1 Zur Messung der Angst: das State-Trait-Anxiety-Inventory (STAI)	181
3.3.1.2 Zur Messung der Depressivi- tät: die Selbststrating-Kurz- form des BECK-Depression- Inventory (BDI)	183
3.3.1.3 Zur Messung der Hilflosigkeit: die Hilflosigkeitsskala (HiS)	187

	Seite
3.3.1.4	Zur Analyse vegetativer Reaktionen: die Beschwerdeliste 188
3.3.1.5	Analogskalen zur Messung der Post-Prüfungsbelastung 189
3.3.2	Die Messung der Strategien 190
3.3.2.1	Strategien I: das Testverhaltensdifferential (TVD) 190
3.3.2.2	Strategien II: die Strategiensammlung (SS) 193
3.3.3	Zur Valididierung durch Fremdbeobachter: das Beisitzerblatt 198
3.3.4	Zum Zielzustand des Kandidaten: die Leistungsdaten 201
3.3.5	Zusammenfassung 202
3.4	Datensätze zur Hypothesengenerierung 203
3.4.1	Zur Wahl der Prüfer durch den Kandidaten 204
3.4.2	Die Kausalattributionierung des Prüfungsergebnisses 206
3.4.3	Die emotionale Reaktion auf die Zensur: die Zensurkonsequenzen 208
3.4.4	Die Variablen zur Erhebung der Prüfungsanmeldemotivation 208
3.5	Datenerhebungs- und Datenverarbeitungsplan 211
3.5.1	Konkrete Planung der Datenerhebung 211
3.5.2	Datenschutzmaßnahmen 215
3.5.3	Darstellung der Konfigurationsfrequenzanalyse und Begründung für ihren Einsatz als Hauptanalysemethode 216
3.5.4	Prinzipielle Planung der Datenanalyse 227

	Seite
4. DARSTELLUNG DER ERGEBNISSE	231
4.1 Deskription der Stichprobe und des natürlichen Designs	231
4.1.1 Die Mortalität der potentiellen und die Rekrutierung der untersuchten Stichprobe	231
4.1.2 Exkurs: Ausführliche und exemplarische Darstellung eines KFA-Ergebnisses	234
4.1.3 Deskription der untersuchten Stichprobe	238
4.1.4 Der tatsächliche Prüfungsverlauf	240
4.1.5 Exkurs: Unterschiede zwischen Teilnehmern und Teilnehmerverweigerern	242
4.1.6 'flaps'	244
4.1.7 Zusammenfassung und Übersicht	245
4.2 Zur internen Validität der erhobenen Daten: Ergebnisse der peripheren Datensätze	247
4.2.1 Der somatische Zustand: die FBL	247
4.2.2 Replikation der Struktur des Arbeitsverhaltens: Analyse des Studiengewohnheitsinventars	250
4.2.3 Replikation des Angstverlaufs	256
4.3 Hypothesenprüfender Teil	271
4.3.1 Analyse der Leistungsdaten (Hypothesen 1 und 2)	271
4.3.2 Analyse der Belastungsmessungen (Hypothesen 6 und 8)	284
4.3.2.1 Analyse der einzelnen Meßinstrumente	284
4.3.2.1.1 Die Hilflosigkeitsskala (HiS)	284
4.3.2.1.2 Die Selbst-Rating-Kurzform des BECK-Depression-Inventary (BDI)	287
4.3.2.1.3 Das State-Trait-Anxiety-Inventary, Form X-1 (STAI)	290
4.3.2.1.4 Analyse der Beschwerdeliste	295

	Seite
4.3.2.1.5 Analyse der Analogskalen zur Erfassung der Nachprüfungsbe- findlichkeit	300
4.3.2.1.6 Analyse der fremdskalierten "Aufregung"	305
4.3.2.2 Agglutination der Datensätze zur Belastungsmessung	306
4.3.2.3 Zusammenfassung und Überblick	318
4.3.3 Analyse der Strategien (Hypothesen 3, 5 und 9)	321
4.3.3.1 Analyse der einzelnen Meß- instrumente	321
4.3.3.1.1 Das Testverhaltensdifferen- tial (TVD)	321
4.3.3.1.2 Exkurs: Analyse der Daten, die von Fremdbeobachtern stammen - das Beisitzerblatt	333
4.3.3.1.3 Analyse der Strategiensamm- lung (SS)	335
4.3.3.1.3.1 Ergebnisse von Faktorenana- lysen des Typs R	336
4.3.3.1.3.2 Selektion von 'Kernstrategie- items'	338
4.3.3.1.3.3 Selektion von 'Trivialitems' und Analyse der Strukturiert- heit	349
4.3.3.2 Zusammenfassung und Überblick	355
4.3.4 Strategien und Belastungen (Hypothese 7, Hypothesentriade zu Vorerfahrungen, Prü- fungsstrategien und Prüfungsbelastungen)	359
4.3.5 Strategien und Leistungsdaten (Hypothese 4, Hypothesentriade zu Vorerfahrungen, Prü- fungsstrategien und Leistungsdaten)	372
4.3.6 Überblick und Zusammenfassung der Ergeb- nisse des hypothesenprüfungen Teils	379
4.4 Hypothesengenerierender Teil	383
4.4.1 Analyse der einzelnen Datensätze	383
4.4.1.1 Die Analyse der Prüferwahlen	383
4.4.1.2 Die Attribuierungsskalen	388
4.4.1.3 Die emotionale Reaktion auf die Zensur	397
4.4.1.4 Die Prüfungsanmeldemotivation	403

	Seite
4.4.2 Agglutination ausgewählter Datensätze	407
5 DISKUSSION	417
5.1 Diskussion der methodischen Grundlagen	417
5.2 Diskussion des hypothesenprüfenden Teils	432
5.3 Diskussion des hypothesengenerierenden Teils	455
6 ZUSAMMENFASSUNG	462
7 ANHANG	467
8 LITERATURVERZEICHNIS	539

LEBENS LAUF

I EINLEITUNG

In dieser Arbeit definieren wir 'Prüfungsangst' als ein Verhaltensmuster, das in einem engen Zusammenhang zur Situation einer Prüfung steht. Dabei beschränken wir uns hier - der Überschaubarkeit wegen - auf Prüfungsangst von Studenten im Zusammenhang mit mündlichen Prüfungen.

Der Begriff der Prüfungsangst beinhaltet zwei Bedeutungskomponenten: die emotionale Reaktion der Angst und den Begriff der Prüfung.

Prüfungen sind in unserer Gesellschaftsform ein notwendiger und allgemein als selbstverständlich erachteter Bestandteil des Studiums an der Institution Universität. Prüfungsforschung stellt sich in der Regel als Teil der Forschungen zur Hochschulelaufbahn und zur Erwachsenensozialisation dar.

Beispielsweise sieht NEUMANN (1933, S. 48) die Prüfung als eine unter vielen Anforderungen des Lebens, "als Abschluß der Vorbereitung auf den Beruf, wozu es - beiläufig - nötig ist, daß den dazu eingesetzten Kommissionen Gelegenheit gegeben wird, sich zu überzeugen, daß man sich Überblick, Wissen und denkmäßige Schulung für den Beruf erworben hat." Der Prüfungsängstliche versucht nach NEUMANN, das Examen in der vorgeschriebenen Form zu umgehen. Überspitzt formuliert heißt das, daß der Prüfungsängstliche sich weigert, seine Hochschulelaufbahn und diesen Teil seiner Sozialisation in der üblichen, beiläufigen Weise abzuschließen.

Nach dieser Interpretationstradition ist Prüfungsangst ein 'Störfaktor' bei der Beurteilung einer Berufsausbildung durch die Prüfung. Aus der Sicht der Institution tritt dieser Störfaktor zufällig auf und wird vom zufällig gerade zu prüfenden Kandidaten in die Prüfung eingebracht. Auf dieser Basis entwickelten sich Forschungsansätze, die Prüfungsangst als ausschließlich intrapsychisches Geschehen interpretieren.

Ein Prüfungsängstlicher, der diese Interpretation auf sein eigenes Leiden und Erleben anwendet, wird seine Angst dann folgerichtig als *privates*, 'persönliches' Problem betrachten.

Dabei ist Prüfungsangst ein erhebliches klinisch-psychologisches Massenproblem. Fragt man nämlich Prüfungskandidaten, ob sie Angst hätten, so erhält man Werte zwischen 66 % (SCHEEF & ZENZ, 1973), 88 % (BROWN & VAN GELDER, 1938), 91 % (PIEHL, 1972) und 96 % (OETTING, 1966). Lediglich HEINERTH (1972) berichtet von 25 % 'Angst'-Angaben, gleichzeitig jedoch von 50 % 'Lampenfieber'.

Wenn wir 'Angst' und 'Lampenfieber' zum Begriff der 'Prüfungsbelastung' zusammenfassen, können wir festhalten, daß die überwiegende Mehrheit der Prüfungskandidaten über Prüfungsbelastungen berichtet.

Auch physiologische Messungen zeigen Prüfungen als Notfallsituationen mit überdeutlichen Streßkomponenten (TURKSON, 1970; OETTING, 1966; BRIDGES, 1973; BROWN & VAN GELDER, 1938; SMITH & WENGER, 1965).

Die Häufigkeit und scheinbare Zwangsläufigkeit der Prüfungsbelastung zwingt den der Praxis verpflichteten Klinischen Psychologen zu einer neuen Perspektive: Er kann sich nicht mehr am Leidensdruck einzelner Betroffener orientieren, er kann seine Interventionsmöglichkeiten nicht mehr auf die Einzeltherapie beschränken, sondern er muß zur Orientierung an prophylaktischen Modellen übergehen.

Die Prophylaxe der Prüfungsangst war das ursprüngliche Erkenntnisinteresse, mit dem wir diese Arbeit begannen. Unsere ursprüngliche Absicht war es, prophylaktische Interventionsmöglichkeiten auszuloten.

Bald stießen wir jedoch auf theoretische Schwierigkeiten, und unsere ursprüngliche Absicht reduzierte sich auf die Suche nach einem geeigneten theoretischen Modell der 'Prüfungsangst'. Wir suchten nach einem Modell, das Gesetzesaussagen enthält, wonach ein - aus der Sicht der Institution Hochschule - ebenso häufiges wie passageres, beiläufiges Ereignis bei der Mehrheit der Betroffenen - bei Zutreffen der Antecedensbedingungen - zwangsläufig starke Emotionen auslöst.

Wir meinen, daß die Theorie der Gelernten Hilflosigkeit (Learned Helplessness, im folgenden daher kurz: LH) von SELIGMAN (1975) diese Bedingungen erfüllt. Gegenstand dieser Arbeit ist daher allein die Übertragung dieser Theorie auf das Gebiet der Prüfungsangst von Studenten in mündlichen Prüfungen.

Dabei wollen wir uns am hierarchischen Modell der Betrachtungsebenen von LEGEWIE & NUSSELT (1975, S. 5f.) orientieren. Die Autoren schlagen vor, verschiedene Organisationsstufen des menschlichen Verhaltens zu unterscheiden und diesen jeweils verschiedene Betrachtungsebenen mit jeweils spezifischen Meßverfahren zuzuordnen. Für den psychophysiologischen Ansatz in der Psychologie kommen sie dazu, vier Betrachtungsebenen zu unterscheiden:

1. Organisch-physiologische Ebene (erfaßbar durch physiologische Registriermethoden),
2. Motorisch-verhaltensmäßige Ebene (erfaßbar z. B. durch direkte Beobachtung),
3. Subjektive Erlebnisebene (erfaßbar z. B. durch Exploration und Fragebögen) und
4. Symbolisch-kommunikative Ebene (erfaßbar z. B. durch Analysen der sozialen Interaktion).
(vgl. LEGEWIE & NUSSELT, 1975, S. 7)

Grundsätzlich gehen wir in dieser Arbeit davon aus, daß die Situation der mündlichen Prüfung aus der Sicht des Kandidaten ein Problem darstellt, das - wie jedes Problem - durch drei Komponenten zu bestimmen ist. Die erste Komponente ist die Ausgangssituation. Deren objektive Charakteristika reichen zur Bestimmung eines Problems jedoch nicht aus, sondern wir benötigen eine zweite Komponente, den Zielzustand. Er unterscheidet sich von der Ausgangssituation und muß aus ihr abgeleitet oder erzeugt werden; sein Erreichen ist mit der Lösung des Pro-

blems identisch. Als dritte Bestimmungskomponente gehört zu einem 'Problem', daß die Überführung der Ausgangssituation in den Zielzustand kein automatischer Vorgang ist, sondern daß erst eine Reihe von Überführungsoperationen gesucht und realisiert werden muß (vgl. KLIX, 1973).

Ein Problem - z. B. eine Prüfung - ist demnach durch drei Komponenten zu bestimmen: Ausgangssituation, Zielzustand und die Menge der verknüpfenden Operationen. Nun gibt es aber (theoretisch) unendlich viele Möglichkeiten, Ausgangssituation und Zielzustand zu verbinden.

Wir nehmen an, daß der Prüfling die theoretisch unendliche Menge von Operationen durch den Einsatz von Strategien reduziert. Als Folge des Einsatzes einer Strategie werden nur noch wenige (endlich viele) Operationen im Problemlösungsprozeß z. B. auf ihre Geeignetheit untersucht.

Ein Subproblem in einer mündlichen Prüfung ist z. B. eine Frage. In dem Moment, in dem der Prüfer die Frage stellt, ist die Ausgangssituation definiert. Gleichzeitig ist aber auch der Zielzustand definiert: die korrekte Antwort. Aufgabe des Kandidaten ist es nun, Ausgangssituation und Zielzustand durch die Ausführung einer oder mehrerer Operationen zu verknüpfen. Dem Prüfling stehen theoretisch unendlich viele Operationen zur Verfügung: er kann lachen, nachdenken, Gegenfragen stellen, die Frage zurückweisen, die korrekte Antwort aussprechen, rülpfen, dem Prüfer in das Ohrchen zwicken usw. Wir haben nun beobachtet, daß Prüfungskandidaten nur einige wenige Verhaltensvarianten ausführen - im Idealfall die korrekte Antwort aussprechen. Andere, durchaus denkbare Operationen werden nicht realisiert. Wir schließen daraus, daß die Kandidaten Strategien einsetzen. Eine sinnvolle Fragestellung kann daher

nicht lauten, ob ein Kandidat Strategien einsetzt, sondern sie lautet, welche Strategien eingesetzt werden.

Der Begriff der Strategie als einer umfassenden, molaren Planung von Operationen umfaßt dabei die Begriffe der 'Erwartung' (= Wahrscheinlichkeit des Eintretens eines Ereignisses) und der 'Taktik' (= konkrete Handlung bei einer Auseinandersetzung).

In der Theorie der Gelernten Hilflosigkeit geht es nun um das Lernen einer bestimmten Strategie (eben der Gelernten Hilflosigkeit) und die Folgen oder Begleiterscheinungen des Einsatzes dieser Strategie. Wir wollen in dieser Arbeit den Nachweis antreten, daß 'Prüfungsangst' nichts anderes ist als die emotionale Begleiterscheinung des Einsatzes dieser Strategie.

Wir werden die Theorie der Gelernten Hilflosigkeit auf das Verhaltensmuster der Prüfungsangst übertragen. Wir gehen zunächst auf die bisherigen Forschungsansätze zu Prüfungen und Prüfungsbelastungen ein. Zu diesem Zweck werden wir die wichtigsten sozialwissenschaftlichen Konzepte zum Begriff der Prüfung darstellen und versuchen, von daher einige Bestimmungskomponenten für das Problem Prüfung abzuleiten. Im psychologischen Teil dieses Kapitels diskutieren wir die Rahmenbedingungen für den Einsatz von Prüfungsstrategien und die bisherigen Forschungsergebnisse zur Symptomatik der 'Prüfungsbelastung'.

In dem darauf folgenden Teil stellen wir die Theorie

der Gelernten Hilflosigkeit und ihre bisherige Entwicklung in drei Phasen dar. Da es bislang kaum kritische - im deutschsprachigen Raum noch gar keine - Darstellungen dieser Theorie gibt, fällt dies im Einzelfall ausführlicher und mit weniger Bezügen zum Thema dieser Arbeit aus, als es zwingend geboten ist. Dem Leser, der weniger an der Theorie der Gelernten Hilflosigkeit interessiert ist, empfehlen wir, nur die Zusammenfassungen am Schluß der einzelnen Abschnitte zu lesen.¹

Im dritten theoretischen Teil vereinigen wir dann die beiden Ausgangspunkte und übertragen die Theorie der Gelernten Hilflosigkeit auf das Prüfungsgeschehen. Dabei legen wir das Modell der Betrachtungsebenen von LEGEWIE & NUSSELT (1975, S. 7) zugrunde. Nun erhebt die Theorie der Gelernten Hilflosigkeit aber den Anspruch, ein Erklärungsmodell für depressives Verhalten zu liefern. Dies bedeutet, daß wir eine Theorie zur Erklärung von 'Depressionen' auf das Phänomen Prüfungs-'Angst' übertragen. Dem Verhältnis von 'Angst' und 'Depression' haben wir deshalb einen Exkurs gewidmet.

¹ Nach Fertigstellung des Manuskripts stießen wir auf einen Sonderband des "Journal of Abnormal Psychology" (1978, Volume 87, Number 1, 198 Seiten) zum Thema "Learned Helplessness as a Model of Depression". Die 17 Artikel dieses Bandes konnten in der vorliegenden Arbeit leider nicht mehr berücksichtigt werden. Nach (flüchtiger) Durchsicht haben wir den Eindruck, daß eine Berücksichtigung zu keiner wesentlichen Änderung geführt hätte.

In weiten Teilen des theoretischen Teils dieser Arbeit verwenden wir den Begriff der 'Strategie' bzw. der 'Prüfungsstrategie' als explikatives Konstrukt, ohne es deskriptiv zu spezifizieren. 'Prüfungsstrategien' in der hier verwendeten Bedeutung sind bislang nicht untersucht worden; insofern ist diese Arbeit auch eine hypothesengenerierende Pilotstudie. Dennoch können wir mit diesem - wenngleich vorerst globalen - Konstrukt Hypothesen ableiten und überprüfen. Wir haben deshalb für das Konstrukt der 'Prüfungsstrategie' eine Mischung aus Hypothesengenerierung und Hypothesenprüfung vorgesehen, die als Analogie zur Logik der Konstruktvalidierung gesehen werden kann. Der Messung von Prüfungsstrategien und ihren Validierungsmöglichkeiten ist das letzte theoretische Kapitel gewidmet.

2 THEORETISCHER TEIL

2.1 Prüfungen und Prüfungsbelastungen

Der Begriff der Prüfungsangst beinhaltet zwei Bedeutungskomponenten: die emotionale Reaktion der Angst und den Begriff der Prüfung.

Die Literatur zum Begriff der Prüfung ist weit verstreut, unterscheidet sich zudem in theoretischer Elaboriertheit, methodischer Genauigkeit und im anthropologisch-gesellschaftstheoretischen Hintergrund. Konsensus besteht wenigstens insofern, als Prüfungen für einen notwendigen Bestandteil des Studiums und der Institution Hochschule gehalten werden. Prüfungsforschung stellt sich häufig als Teil der Forschung zur Hochschullaufbahn, zur Hochschulsozialisation dar. Mit unterschiedlicher Gewichtung unterscheiden die im folgenden referierten Ansätze eine

- statuszuweisende,
- sozialisationsabschließende,
- studiensteuernde und
- sozialselegierende (bewertende)

Funktion von Prüfungen.

Wenn wir davon ausgehen, daß Prüfungen eine Problemsituation darstellen, mit der sich ein Student problemlösend auseinanderzusetzen hat, dann erscheint es sinnvoll, zunächst die wichtigsten theoretischen Ansätze zur Interpretation dieser Problemsituation darzustellen.

2.1.1 Die Funktionen von Prüfungen aus sozial- wissenschaftlicher Sicht

"Über ein Drittel der Forschungsliteratur, die Aspekte des Hochschulwesens behandelt, beschäftigt sich mit den 'subjektiven Entwicklungsprozessen' in Verbindung mit akademischen Institutionen eines Landes. Je nach fachspezifischer Perspektive werden sie als Sozialisation, Persönlichkeitsentwicklung, Enkulturation, Akkulturation, Entscheidungsprozesse und Lernprozesse konzeptualisiert." (NITSCH, 1973, S. 375)

NITSCH unterscheidet siebzehn Forschungsdisziplinen, die an Forschungen über subjektive Prozesse im Kontext akademischer Institutionen beteiligt sind. Größere Forschungsprojekte kumulieren dabei mehrere dieser fachspezifischen Ansätze. Einige davon erleben z. Z. eine Forschungsexplosion, andere wurden bislang völlig vernachlässigt oder nur sporadisch bearbeitet.

Die größte Wachstumsrate haben z. Z. die sozialstatistischen und psychometrisch-ökologischen Erhebungen. Sie dienen der Prognose von Studienerfolg und -mißerfolg, der Entdeckung von Begabungspotentialen, der Erklärung von Leistungsvarianzen. Sie dienen z. T. als Grundlage für Selektionsmaßnahmen (z. B. für weiterführende Kurse) und für Personalsteuervorhaben (vgl. z. B. KÜHN & TANNENBERGER, 1971; HACKER, KÖHLER & KÜHN, 1964; SCHALLBERGER, 1974; KATH, KÖHLER & REICHWEIN, 1967). Auch Korrelationsstudien über den Zusammenhang von Persönlichkeitsvariablen und erhaltenen Zensuren zählen zu dieser Gruppe von Forschungsansätzen (vgl. z. B. HECKHAUSEN, 1977; TURKSON, 1970, SCHEER & ZENZ, 1973).

Völlig vernachlässigt sind nach NITSCH rechtssoziologische, kulturen- und kulturanthropologische, soziolinguistische und sozial-psychophysiologische Ansätze.

"Die Kontroversen und Differenzierungen zwischen theoretischen Traditionen und Schulen lassen sich insbesondere an der Genesis, der erkenntnislogischen Struktur und der heuristischen Relevanz der zahlreichen Konzepte und Paradigmen der Forschungen über subjektive Entwicklungsprozesse... konkretisieren." (NITSCH, 1973, S. 382)

"Je nach dem Grad der semantischen Differenzierung lassen sich etwa 150 bis 200 Paradigmen und Konzepte isolieren, die entweder der Analyse, der Phänomenologie, der Kritik und der Steuerung subjektiver Prozesse und Faktoren in Verbindung mit akademischen Institutionen dienen." (NITSCH, 1973, S. 384)

Die verschiedenen Konzepte reflektieren unterschiedliche Erkenntnisinteressen, Realitätsperspektiven, Operationalisierungen und Instrumentarien verschiedener am Hochschulprozeß beteiligter Gruppen: der Forschenden, der Lehrenden, der Prüfenden, der Administrierenden, der Politiker, zuweilen sogar der betroffenen Studenten. Wir wollen versuchen, die verschiedenen Konzepte unter vier übergreifenden Richtungen zu subsumieren (vgl. NITSCH, 1973).

2.1.1.1 Anthropologische Konzepte

Anthropologische Konzepte aus der Kultur- und Sozialanthropologie liefern die ältesten Paradigmen für Studien über die Entwicklung der Subjekte in akademischen Institutionen. Als Basiskonzepte dienen 'Akkulturation' und 'Enkulturation', d. h.

der Erwerb von Kulturgütern von Studenten über Lehrer (Akkulturation) oder durch andere Studenten (Enkulturation). Hierzu zählt auch das Paradigma von College-Subkulturen, z. B. die Vorstellung, Studenten bildeten eine relativ homogene und nach außen hin abgeschottete Population, die z. B. dadurch entsteht, daß die Studenten aus ihrer gewohnten familiären und subkulturellen Heimatumgebung herausgerissen werden und in eine neue Stadt/Klasse/Religion/Kultur/Milieu hineinwachsen (vgl. hierzu KÜHN, 1970, 1971; KÄPPLER, 1977). Dieser Vorgang wird besonders bei den amerikanischen 'residential colleges' deutlich, wo die Studenten praktisch auf dem Campus wohnen.

Hierzu zählen auch die Konzepte des Initiationsritus' und die analytische Interpretation der Universität als statuszuweisender Institution. Die Prüfungsangst wird von KVALE (1972), ZENZ & SCHEER, (1973), SCHEER & ZENZ (1973), FLUGEL (1939) und STENGEL (1936, 1938) als Korrelat des Initiationsritus' der Prüfung gesehen, in dessen Verlauf die Erwachsenengesellschaft dem Heranwachsenden seinen zukünftigen Status verleiht.

Weitere Konzepte sind die des idealen Studenten (COX, 1966), des College-Slang, der Fachsprache, des Kulturschocks (insbesondere bei Ausländern), des Realitätsschocks (bei Studienanfängern, Hochschulwechslern und Examenskandidaten), evtl. auch eine reduktionistische Version des MARXschen Entfremdungskonzepts (vgl. KÄPPLER, 1977; TURKSON, 1970; KVALE, 1972).

Unsere heutigen akademischen (mündlichen) Prüfungen werden demnach - wie schon erwähnt - als Analogie oder Äquivalent des Reiferituals primitiver Völker gesehen. Die Funktion dieses Rituals ist die Aufnahme des Heranwachsenden in die Gruppe der Erwachsenen bzw. der Wechsel von einem Minoritäts- zu einem Majoritätsstatus. Die von uns zitierten Veröffentlichungen stellen den Ritus so dar, daß der junge Erwachsene nach diesem einmaligen Akt der Initiation an allen Privilegien seines neuen Status teilhaben darf. Dadurch tritt er jedoch in Konkurrenz zu den älteren (und weniger leistungsfähigen) Erwachsenen. Diese sind deshalb daran interessiert, den Initiationsritus möglichst schwierig und belastend geraten zu lassen. Ein Reiferitual ohne Angst (eine Prüfung ohne Prüfungsangst) verfehlt aus anthropologischer Sicht ihren Zweck.

Die vorgestellte kulturanthropologische Interpretation bleibt jedoch einseitig. Sie geht davon aus, daß es für die Gruppe der Erwachsenen einen endlichen Nahrungskuchen gibt. An diesen will der Prüfling heran. Die Erwachsenen teilen aber nicht so gern und machen es ihm deshalb schwer. Dabei wird übersehen, daß selbst in einer primitiven Gesellschaft der Prüfling nach dem Ritus seine volle Arbeitskraft zur Verfügung stellen kann und zugleich durch den Ritus eventuelle Privilegien seines alten Status' verliert (z. B. das relativ sorglose und beschützte Leben).

Die Faszination dieser Interpretation liegt unseres Erachtens weniger in dem Erklärungswert, sondern mehr in dem ironischen Verweis auf tatsächliche,

rituelle, unbegründete Aspekte von Prüfungen, insbesondere von mündlichen Prüfungen.

Vereinzelte finden sich an den Universitäten immer noch Prüfungen, in denen besondere Kleidung - z. B. schwarze Anzüge für die männlichen Prüfungskandidaten - getragen wird. Eine (meist) anonyme Institution - das 'Prüfungsamt' - legt Ort und Zeit der Prüfung fest; bei einzelnen Fächern werden sogar die Prüfungsthemen nicht nur dem Prüfling, sondern auch dem Prüfer vorgegeben. Dem Prüfungskandidaten wird meist nicht gestattet, Ort und Zeit der Prüfung zu beeinflussen; jedes Abweichen wird sanktioniert, es sei denn, der Kandidat wählt einen rituellen Ausweg: er wird krank. NEUMANN (1933) sieht sich z. B. gehalten, die 'Flucht in die Krankheit' explizit als Vermeidungsverhalten aufzuführen. Vereinzelte wird sogar überprüft, ob Krankheit oder 'Krankheit' vorliegt, indem erst das Attest eines vereidigten Amtsarztes von der Prüfungspflicht entbindet. In den Prüfungen selbst wird ein 'individueller Leistungsnachweis' selbst dann verlangt, wenn diese Leistung im Berufsleben nur im Team erbracht wird. Und wenn ein Prüfungskandidat für all dies wenig Verständnis zeigt, dann erlebt er das Ritual des Arguments der verwaltungstechnischen oder juristischen Notwendigkeit.

Die von den kultur- und sozialanthropologischen Standpunkten abgeleitete statuszuweisende Funktion von Prüfungen ist auch Gegenstand der Gruppe der soziologischen Konzepte, auf die wir im folgenden eingehen.

2.1.1.2 Soziologische, systemfunktionalistische Konzepte

Soziologische Konzepte stehen in der Tradition von

"... Studien über soziale Systeme und ihre Strukturelemente Stratifikation, Mobilität, Status-Allokation einerseits und über 'Sekundärsozialisation' oder Erwachsenensozialisation andererseits. ... Sekundärsozialisation wird als das Lernen der - unterschiedlich theoretisch gefaßten - Verbindung von individueller 'Motivation' oder Triebimpulsen mit sozialen Rollen, die von bestimmten engeren Gruppen oder Kategorien von Personen in einer Gesellschaft erfüllt werden und/oder erfüllt werden sollen, verstanden, in Abgrenzung von 'Primärsozialisation' oder 'Sozialisation', dem Erwerb der für ein ganzes Gesellschaftssystem (in schichtspezifischer Varianz) geltenden Prinzipien der Verbindung individueller Triebimpulse mit elementaren allgemeinen sozialen Normen dieses Gesellschaftssystems." (NITSCH, 1973, S. 396).

Hierzu zählen alle Ansätze, die die Rekrutierungsfunktion von Prüfungen betonen oder beinhalten. KVALE (1972), SCHEER & ZENZ (1972), insbesondere auch KÄPPLER (1977), sehen Prüfungen als ein Instrument oder als Bestandteil eines Instruments, mit dem die Gesellschaft ihre intellektuelle Elite selektiert bzw. mit dem die bürgerliche Elite überwiegend sich selbst rekrutiert. Dies geschieht durch Selektion in der Prüfung und anschließende Statusverleihung.

Verschiedene Schwerpunkte in der Theoriebildung können unterschieden werden (NITSCH, 1973, S. 397f.), wenngleich eine funktionalistisch-kybernetische (systemfunktionalistische) Betrachtungsweise vorzuherrschen scheint. Für die Prüfun-

gen im engeren Sinne relevant sind z. B. das Modell der 'pattern variables', wonach bis zur Prüfung bestimmte Verhaltensmuster ausgebildet und dann geprüft werden (universalism vs. particularism, quality vs. performance, diffuseness vs. specificity, affectivity vs. affective neutrality). Wichtig ist auch das Konzept des 'role-making', "... ein Prozeß, der in kulturanthropologischer Sicht als Überlagerung von (beruflicher) Kulturgenese, Enkulturation, Akkulturation und interkultureller Austauschprozesse zu betrachten wäre, in berufssoziologischer Sicht dagegen als Überlagerung von (objektiver) Professionalisierung, (subjektiver) Personalisation (als Sekundärsozialisation), Integration in eine komplexe professionelle Organisation und inter-professioneller (horizontaler) Kooperation/Konsultation/Kommunikation zu formulieren wäre;" (NITSCH, 1973, S. 400).

"Der dominierende systemfunktionalistische Trend in den soziologischen Analysen der Entwicklung von Studenten, Akademikern und Wissenschaftlern kommt in vielen Konzepten zum Ausdruck, die Vorstellungen von einer relativ starken Abgeschlossenheit, Autonomie, Stabilität, Konfliktlosigkeit und Homogenität von Institutionen und Organisationen bekräftigen, in die Individuen zur formalisiert zielgerichteten Veränderung oder Erweiterung ihrer Persönlichkeitsmerkmale, Einstellungen, Werthaltungen, Kenntnisse oder Fertigkeiten temporär eingegliedert werden." (NITSCH, 1973, S. 402)

Viele der genannten Konzepte stammen aus Forschungen über geschlossene Anstalten (Gefängnisse, Ferienlager, Militäreinheiten) und werden statt als Analogie als Äquivalent in die Analyse von Entwicklungsprozessen (und deren Überprüfung in anderen Institutionen) übernommen.

"Der exzeptionelle, stark formalisierte und manipulationsbedürftige Charakter solcher geschlossener Anstalten sowie ihre Verbindung mit politischen und administrativen Aufgaben von staatlichen und kirchlichen Behörden haben offenbar mehr soziologische Forschungs- und Beratungstätigkeit angereizt als das Studium von primär subjektiven Prozessen in Verbindung mit einzelnen oder mehreren offenen, öffentlichen, inhomogenen, konflikthaften oder weniger formalisierten Institutionen (...), deren Struktur den meisten Hochschulsystemen, vor allem den 'non-residential universities' oder den 'commutor colleges' (Pendler-Fachhochschulen) vergleichbar wäre, und die - wie diese - schwieriger durch 'social engineering' zu steuern und zu verändern sind...

In den Studien und social-engineering-Strategien über residential colleges und studentische Wohnheime (...) oder über Sozialisation und Ausbildungsgänge im Militär (...) lassen sich die aus der Erforschung und Kontrolle geschlossener Anstalten gewonnenen Paradigmen dagegen direkt übertragen, vorausgesetzt, daß sich die Geschlossenheit und Kontrolle - zum Teil mit Hilfe der diese Konzepte anwendenden Forscher - erfolgreich erzwingen lassen." (NITSCH, 1973, S. 403f.)

Das wichtigste, gleichzeitig formale und einfache Grundkonzept innerhalb der funktionalistisch-kybernetischen Soziologie ist das der 'Sozialisationsagentur', wonach gesellschaftlich legitimierte Agenten - die Prüfer - den erfolgreichen Verlauf oder Abschluß der Sozialisation überprüfen.

NITSCH geht ausführlich auf diesen Ansatz ein, dessen Erforscher

"... unter dem Einfluß der strukturell-funktionalistischen Primär-Sozialisierungstheorie von PARSONS und BALES anzunehmen scheinen, daß auch zu den Grundbedingungen jeder erfolgreichen Sekundärsozialisation, zumindest soweit sie sich auf komplexe, kulturelle Normen bezieht, gehört, daß sozialisie-

rende Personen, die mit überlegenen Machtmitteln ausgestattet sind, zusammen mit den 'schwächeren', von ihnen abhängigen, zu sozialisierenden Personen ein institutionalisiertes soziales System bilden und daß die überlegenen 'Sozialisierer' durch Internalisierung fest in das kulturelle Wertsystem integriert sind, das das formale Ziel der Sozialisation darstellt. Hochschulen und akademische Berufsorganisationen sind in dieser Sicht absichtsvoll eingerichtete und formal organisierte Subsysteme oder Subgesellschaften, zusammengesetzt aus Handlungen von Individuen und Konstellationen von Rollen mit dem Ziel der Problemlösung und der Erweiterung des Rationalen zu Lasten der (naturwüchsig)-institutionalen Komponenten sozialer Interaktion (...)" (NITSCH, 1973, S. 405)

Dieses Grundmodell ist auf verschiedene Weise weiterentwickelt und differenziert worden. NITSCH erwähnt z. B. das Konzept der psychischen Adaptation an das soziale Subsystem, das Konzept der 'hidden or invisible curricula', verschiedene Rollenkonzepte, darunter das Konzept der Studentenrolle als einer sich selbst liquidierenden Rolle u. a.

Die Inhalte der Sozialisation können einerseits faktorenanalytisch in Form von 'Hochschulklimate syndromen' oder in Form strukturell-funktionalistischer Taxonomien (BLOOM et al., 1956) bestimmt werden.

OLESON & WHITTAKER (1966, zit. n. NITSCH, 1973) kritisieren aus der Position des symbolischen Interaktionismus, daß die systemfunktionalistischen Ansätze sich fast alle zum Bild der Sozialisation als 'Beraubung' oder 'Entkleidung' und des Studenten als eines 'Neugeborenen' zusammenfinden: Der Student wird beim Eintritt in eine

neue institutionalisierte Entwicklungsphase seiner bisherigen Eigenschaften beraubt und in einer totalen Enkulturation gleichsam neugeboren.

Dadurch aber können drei systemfunktionalistische Paradigmen in einem kleinsten gemeinsamen Nenner zusammengefaßt werden:

- Konzepte aus der Tradition soziologischer oder sozialpsychologischer Studien zur kindlichen Entwicklung,
- Konzepte aus der Tradition von Forschungen über geschlossene Anstalten und
- Konzepte aus der Tradition von Forschungen über die Vermittlung von Werten, Einstellungen und Motiven durch Massenkommunikationsmittel.

Systemfunktionalistische Analyseansätze zeichnen sich zusammenfassend dadurch aus, daß eine Institution mit klaren Regeln als unveränderbar vorgegeben wird. Der Student als der relativ Ohnmächtige wird im Verlauf seines Studiums - ob er will oder nicht - sozialisiert und letztlich - was er gar nicht verhindern kann - aufgesaugt. Prüfungen können dabei zweierlei leisten, sie steuern die Sozialisation und bilden deren bewertenden Abschluß.

Die Studiensteuerung durch Prüfungen ist für jeden Studenten ein vertrautes Phänomen. Zum einen wird in den meisten Prüfungsordnungen das Prüfungsergebnis als Strafreiz konzipiert, mit dem durch negative Verstärkung das Arbeitsverhalten angeregt wird.

(Eine Förderung des Arbeitsverhaltens ist auch durch Belohnung - eine gute Zensur - möglich. In Prüfungsordnungen wird diese Möglichkeit meist nicht aufgegriffen. Dafür werden ausführliche Bestimmungen für den Fall formuliert, daß ein Kandidat die Prüfung nicht besteht.)

Zum anderen orientiert sich der Student mit dem zeitlichen Heranrücken der Prüfung mehr und mehr an dem geforderten Prüfungsstoff und nutzt die - z. B. durch eine fehlende Studienordnung - gegebenen Freiräume weniger und weniger aus. In der Zeit der unmittelbaren Prüfungsvorbereitung bestimmt der Prüfungsstoff dann vollständig die Arbeit des Kandidaten.

Latent kann dabei ein verborgenes Curriculum wirksam sein, ein Konzept, das die expliziten kognitiven Prüfungslernziele um nichtkognitive Teile erweitert. Gegenstand der Prüfung sind demnach nicht nur die definierten Stoffgebiete, sondern auch ein bestimmtes Rollenverhalten, bestimmte Werte, eine bestimmte sprachliche Diktion, insbesondere bestimmte Affekte. Diese gehen dann in die Bewertung der Prüfungsleistung ein.

Systemfunktionalistische Ansätze haben sich bei der Analyse der Prüfungsangst meist einseitig und zumeist stillschweigend entfaltet. Dabei können drei komplementäre Richtungen unterschieden werden:

1. Prüfungsangst als respondente Reaktion (Phobie), (vgl. z. B. GARLINGTON & COTLER, 1968)
2. Prüfungsangst als interferierendes Reaktionsmuster (vgl. z. B. MANDLER & SARASON, 1952) und

3. Prüfungsangst als unangemessenes Reaktionsmuster (vgl. z. B. WINE, 1971).

Für jede einzelne dieser Richtungen liegen zahlreiche - hier im einzelnen nicht aufgeführte - empirische Untersuchungen vor. Gemeinsam ist diesen Ansätzen, daß die Institution über die Prüfungsanforderungen als unveränderbar vorgegeben und die Prüfungsangst als 'Störvariable' bei der Überprüfung des Sozialisationserfolges konzipiert wird. Der beschränkte Integrationswert dieser Ansätze setzt sich in der Auffassung fort, daß Prüfungsangst - da sie institutionell nicht verankert sein kann - eine 'individuelle Reaktion' sein müsse, deren Genese in der individuellen Vorgeschichte zu liegen habe. Prüfungsangst wird so zum intrapsychischen Phänomen, das nur zufällig mit der Situation Prüfung assoziiert ist.

Konsequent wird die Genese der Prüfungsangst als Entwicklung von (zufälligen) klassischen Konditionierungen oder als Folge von (individuellen) Persönlichkeitsfaktoren dargestellt.

Damit kommen wir zur Schlußfolgerung, daß einerseits das systemfunktionalistische Konzept einer totalen Institution (ungerechtfertigt) übernommen wird, andererseits zahlreiche empirische Arbeiten aber gerade dieses Konzept (unzweifelhaft) in Beziehung zur Prüfungsangst gesetzt haben. Wenn wir ein Gedankenexperiment wagen und diese Beziehung umkehren, dann würden wir eine offene Institution mit veränderbaren Regeln in Beziehung zur 'Freiheit von Prüfungsangst' setzen müssen.

Damit kämen wir zur Variation der Variablen 'Totalität der Institution', wie sie ansatzweise in den symbolisch-interaktionistischen Ansätzen auftaucht.

2.1.1.3 Symbolisch-interaktionistische Ansätze

Die symbolisch-interaktionistischen Ansätze sind

"... an der komplexen Dreiecksbeziehung von Institution (Organisation und Bezugspersonen), Subjekt (ego, self) und der Totalität von parallel ablaufenden Entwicklungs-, Rollen- und Statussozialisationsprozessen interessiert, vorzüglich unter dem Gesichtspunkt der Entstehung und prekären Aufrechterhaltung persönlicher und sozialer Identität des Subjekts (...)" (NITSCH, 1973, S. 420)

Das wichtigste Konzept ist dabei das des 'psyching out', "... des 'Herausführens' von Signalen des Rollenverhaltens signifikanter Anderer unter Bedingungen der 'Unsicherheit' oder 'Ambiguität' oder der (manifesten) 'Widersprüchlichkeit' der Rollen-Anforderungen und der Rollenkonzeption der signifikant Anderen ..." (NITSCH, 1973, S. 433f.)

"Dieses Paradigma ist den offenen und konflikthafte Rollen-Definitionen in vielen höheren Bildungsprozessen und akademischen Arbeitsprozessen besonders angemessen. Ähnlich wie das Paradigma der 'Rollenmaskerade' wirkt es aber in seiner phänomenologischen Begrenztheit als suggestive Bekräftigung eines Stabilisierungs- und Verdrängungsmechanismus' in herrschaftlich organisierten Sozialisations- und Rollen-Konstellationen. Es beschreibt nur 'einführend' die Strategie, die einen Zusammenbruch der allseitigen Rollenmaskerade verhindert, und es setzt die naturwüchsige Unterwerfung des Sozialisanden unter tendenziell widersprüchliche, unklare, unverstandene, ... Rollen-Anforderungen

voraus, die als unbegriffene, den Subjekten fremde, phantomhaft-normale Macht diese in ihren Bann schlagen." (NITSCH, 1973, S. 434)

NITSCH kritisiert im weiteren, daß in diesen Theorien Konzepte wie Rollen-Widerwillen, Rollen-Ab-lehnung u. ä. nur vereinzelt und wenig differenziert auftauchen.

Dabei sind solche Überlegungen im Hinblick auf Prüfungen durchaus notwendig, wo Prüfer und Kandidat zum Rollenspiel 'Prüfung' zusammentreffen. Die Weigerung eines der beiden, die für ihn vorgesehene Rolle zu übernehmen (bzw. in abgeschwächter Form der Widerwillen gegen die Rollenübernahme) müßte das - aus systemfunktionalistischer Sicht - institutionalisierte Rollenspiel in erheblicher Form stören. Überlegungen aus psychologischer Sicht, was z. B. die beiden dann empfinden, oder aus soziologischer Sicht, welche Korrekturmechanismen die Institution dann vorgesehen hat, fehlen vollkommen.

Für den Bereich der akademischen Sozialisation haben BECKER et al. (1958, 1961, 1968, zit. n. NITSCH, 1973) ein umfassenderes heuristisches Modell herausgearbeitet: das Modell der Perspektiven. Es ist definiert durch die geteilten Einstellungen von Personen zur Situationsdefinition, ihre dementsprechenden Aktivitäten und die Kriterien der Beurteilung, wobei eine grundsätzliche Dichotomie zwischen langfristiger und kurzfristiger Perspektive angenommen wird. Die sogenannte 'Grade-Point-Average-Perspective (GPA)' beschreibt ein studentisches Verhalten,

bei dem ein Kompromiß aus den durch das Studium erreichbaren Lebenszielen und dem Interesse an einem angenehmen und wenig anstrengenden Studentenleben wirksam ist. Die wichtigsten Komponenten aus der Sicht der Studenten sind 'psyching out' und 'fronting' (= Rollenmaskerade) als aktive, aber submissive Strategien der Auseinandersetzung mit der Institution, 'discount' als emotionale Abwehrhaltung gegenüber einer intensiv dominanten Einflußquelle, zyklenhafte Subkulturbildungen, in Grenzsituationen Mogeln und zeitweiliger Studienabbruch. Auch hier dominiert freilich die Vorstellung des Studenten als eines Insassen einer Sozialisationsanstalt, der aus funktionalen Gründen temporär unterworfen ist.

Leider sind BECKERS Arbeiten nicht mehr verfügbar; sie zeigen jedoch Perspektiven für die Analyse der Prüfungssituation und der LH-Experimente.

2.1.1.4 Ansätze der Kritischen Theorie

Einen wichtigen Beitrag zum Prüfungswesen und damit auch zur Prüfungsangst bieten die Autoren, die der Kritischen Theorie nahestehen. Exemplarisch, weil als einziger explizit auf unser Thema bezogen, sei hier der Ansatz von KVALE (1972) vorgestellt.

KVALE sieht die Prüfung als Teil eines Produktionsprozesses, in dem die Ware 'akademische Arbeitskraft' hergestellt wird. Wie bei jeder Ware kann ein Gebrauchswert (wozu wird die Ware gebraucht?)

und ein Tauschwert (was bekommt man dafür?) unterschieden werden. Der Gebrauchswert der akademischen Arbeitskraft liegt z. B. darin, daß Brücken gebaut werden, Kranke geheilt werden, Kindern etwas gelehrt wird. Der Tauschwert der akademischen Arbeitskraft findet einen unmittelbaren Ausdruck in dem später meist überdurchschnittlichen Gehalt. KVALE sieht den Tauschwert in erster Linie als durch die Länge der Ausbildung bestimmt an, verweist aber gleichzeitig auf dessen Marktabhängigkeit.

Ein Überangebot an akademischer Arbeitskraft läßt den Marktwert sinken - die Berufsanfänger müssen niedrigere Anfangsgehälter in Kauf nehmen -, umgekehrt steigt der Marktwert bei einem Unterangebot. Bestimmte Marktmechanismen stehen zur Verfügung, um Angebot und Nachfrage zu regulieren.

Neben der Länge der Ausbildung bestimmt die Prüfungszensur den Tauschwert eines Examsens.

"Auf dem Warenmarkt werden qualitativ verschiedene Gebrauchswerte vergleichbar gemacht durch Quantifizierung ihrer Tauschwerte auf einer Dimension - der des Geldes. Auf dem akademischen Warenmarkt werden qualitativ verschiedene Leistungen unter Hinweis auf Ausbildungsdauer, den akademischen Rang des Abschlußexamens und dessen Noten ebenso vergleichbar gemacht." (KVALE, 1972, S. 175)

Die Reduktion der akademischen Arbeitskraft auf ein eindimensionales Kontinuum - die Zensurenskala - ist jedoch nur ein Parameter von Prüfungen. Gleichzeitig wird nach KVALE vom Prüfungskandidaten erwartet, den Warencharakter seiner Ausbildung zu akzeptieren und zu internalisieren.

Der Prüfungskandidat muß sich im Verlauf seines Studiums nämlich einige Merkmale des Prüfers - Wissen, Kenntnisstand, Werthaltungen - aneignen, er muß sich Merkmale der Mündigkeit zulegen. Der Prüfer bestimmt jedoch aus seiner Machtposition heraus den Tauschwert der Ausbildung durch die Zensur; dem Kandidaten verbleibt die Rolle des unmündigen Abhängigen. Dieser potentielle Konflikt - Unmündigkeit vs. Mündigkeit - wird während der Prüfungssituation dadurch gelöst, daß er die quantifizierende Bewertung seiner Person durch Vorgesetzte internalisiert und übernimmt.

Dem internalisierten Gebrauchswert entsprechen Arbeitsfreude, die Überzeugung der eigenen Tüchtigkeit, die fachliche Identität: die Aneignung von Fachwissen wird zu einer Art Funktionslust. Dem internalisierten Tauschwert entsprechen hingegen Konkurrenzdenken, Karrieresucht und die Fähigkeit, sich mit dem Ellenbogen durchzusetzen: Kenntnisse werden erworben, um über gute Noten zu gutem Geld zu kommen. Diese extrinsisch motivierende Funktion des Zensurensystems führt dazu, daß die Note zum Fetisch und vom Arbeitsprozeß losgelöst kogniziert wird. Das Resultat ist Entfremdung und eine Verdinglichung der sozialen Beziehungen zu anderen Subjekten: die 'Akkulturation' in das kapitalistische System hat stattgefunden.

Die Internalisierung des potentiell konfliktträchtigen 'Herr-Knecht-Verhältnisses' in der Prüfung wird durch die zeitliche Lösung des Konflikts erleichtert: nach Durchlaufen des Initiationsritus' wird der Status des Herren verliehen.

KVALE übernimmt Elemente aus allen von NITSCH (1973) unterschiedenen Forschungsansätzen. Er greift die Idee des Initiationsritus auf, diskutiert die statusverleihende Funktion von Prüfungen, betont den kulturanthropologischen Aspekt der Akkulturation, greift den existenzsoziologischen Aspekt mit dem Begriff der Entfremdung auf und folgt dem funktionalstrukturellen Ansatz mit der Interpretation der Universität als Warenproduktionsstätte, die ökonomischen Gesetzen unterliegt.

Der strukturell-funktionalistische Ansatz mit seiner inhärenten Vorliebe für die Betrachtung geschlossener Systeme wird von KVALE ausgeführt in dem Konzept der rational strukturierten Bildungsfabrik Universität, die den Steuerungsvorgängen eines dynamischen, gleichwohl in sich geschlossenen kapitalistischen Systems gehorcht. Der Prüfling wird dabei als ein Subjekt konzipiert, das sich der übermächtigen Sozialisationsagentur und der dort geforderten Sozialisation nur ganz entziehen kann (z. B. durch Studienabbruch) oder aber sich letztlich total den Erziehungsmechanismen ausliefert. KVALE ist nicht so konsequent, den Begriff der Fabrik wörtlich zu nehmen, nämlich als einer Produktionsstätte, die man morgens aufsucht und abends wieder verläßt.

Wesentliche Elemente der Erwachsenensozialisation müßten sich dann außerhalb der Universität vollziehen. Diese Überlegung greift KVALE jedoch nicht auf.

Auch wenn wir KVALE zustimmen, daß die Universität auch der Ausbildung von Arbeitskräften dient, so

können wir der impliziten Gleichsetzung von Tauschwert und Zensureniveau nicht unbedingt folgen. Zwar gibt es Indizien dafür, daß das Ausbildungsniveau und auch die Höhe der Zensur ihren Niederschlag im Einkommen finden (vgl. AGER, 1966/1967), der Nachweis einer linearen Beziehung fehlt jedoch unseres Wissens. Indizien finden wir in Zeitschriften wie ERZIEHUNG UND WISSENSCHAFT (1977, 4, S. 3), daß in einer Konferenz des Bundesinnenministeriums von einer entsprechenden Reaktion der freien Wirtschaft durch Vertreter des Wirtschaftsrats berichtet worden sein soll.

Vermutlich ist es angemessener, in der Zensuren-skala einen oder mehrere Cut-off-Punkte (vgl. CRONBACH & GLENER, 1965) anzunehmen. Danach müßte der Prüfungskandidat eine bestimmte Mindestzensur erreichen, um bei der Besetzung einer bestimmten Stelle oder der Festlegung eines bestimmten Gehalts berücksichtigt zu werden. Die Annahme eines Cut-off-Punktes beinhaltet, daß der potentielle Arbeitgeber unter der Menge der Bewerber zunächst eine Vorentscheidung nach der Zensur trifft: Alle Bewerber, die eine schlechtere als die Cut-off-Zensur erhalten haben, werden aussortiert. Für die verbleibenden Bewerber werden dann andere Kriterien verwendet.

Die Ausführungen von KVALE beleuchten noch einen wichtigen Aspekt des Prüfungswesens: den funktionalen Stellenwert einer Prüfung. Auch wenn in jeder Prüfung eine Zensur vergeben wird, müssen wir doch unterscheiden zwischen Prüfungen, in denen die Zensur einen direkten Bezug zum Tauschwert haben könnte - wie z. B. in Abschlußprüfungen wie der Hauptdiplom-

prüfung für Psychologen - und zwischen Prüfungen, in denen das Erreichen eines Cut-off-Punktes den Erfolg sichert - im einfachsten Fall muß die Prüfung lediglich bestanden werden wie in den Diplomvorprüfungen für Psychologen.

Ein Vordiplomand in Psychologie 'profitiert' in keiner Form von sehr guten Leistungen, wenn man ihn mit einem Kandidaten vergleicht, der durchschnittliche Ergebnisse erreichte. Sowohl Prüfungsordnung wie Prüfungssystem verlangen lediglich das Bestehen der Prüfung, damit der Kandidat sein Studium fortsetzen darf. Der Tauschwert der Zensur ist nahezu Null. Der Gebrauchswert ist allenfalls aus der Sicht der Institution dahingehend einzuschätzen, daß die weitere Ausbildung durch nachgewiesene Grundkenntnisse erleichtert wird.

Die Kalkulation für Hauptdiplomanden sieht anders aus. Der Gebrauchswert seiner Prüfung ist durch Studien- und Prüfungsordnung vorgezeichnet. Für ihn geht es um eine Maximierung des Tauschwertes, um eine möglichst gute Zensur. Es steht also zu erwarten, daß der Hauptdiplomand in stärkerem Ausmaß als ein Vordiplomand ein Verhalten realisiert, das den Tauschwert erhöht und mit einer guten Zensur honoriert wird. Daraus folgt direkt die Hypothese 1, daß Hauptdiplomanden im Schnitt bessere Zensuren erhalten als Vordiplomanden. Gleichzeitig folgt daraus, daß Vordiplomanden und Hauptdiplomanden bzw. unterdurchschnittlich und überdurchschnittlich Zensierte sich durch konkret angebbare Verhaltensweisen und/oder Strategien unterscheiden (Hypothese 2).

Die Konzeption der Universität als einer nach (teilweise) ökonomischen Gesichtspunkten organisierten Institution, in der Prüfungen die Studiensteuerung, die Selektion und die Initiation übernehmen, läßt die Folgerung naheliegen, daß vor allem der gesellschaftliche Selektionsauftrag den Gesichtspunkten ökonomischer Effizienz unterliegt.

Daraus folgt, daß Gesichtspunkte der Kostenminimierung und Nutzenmaximierung eine möglichst frühzeitige Selektion erfordern, ein möglichst frühzeitiges Aussortieren 'ungeeigneter' Studenten.

Für den oben angeschnittenen funktionalen Unterschied zwischen Prüfungen bedeutet dies, daß Vordiplomanden schärferen Selektionsmechanismen unterworfen werden als Hauptdiplomanden. Im strengen Sinne ist damit nichts über die Durchschnittszensur gesagt, sondern lediglich über die Durchfallquote. Unsere Hypothese lautet daher, daß mehr Vordiplomanden die Prüfungen nicht bestehen als Hauptdiplomanden (Hypothese 3).

2.1.1.5 Zusammenfassung und Diskussion

Der Begriff der Prüfungsangst beinhaltet zwei Bedeutungskomponenten: die emotionale Reaktion der Angst und den Begriff der Prüfung.

Die vorgestellten Ansätze legitimieren die Durchführung von Prüfungen mit

- dem Schutz der Erwachsenen (Initiationsritus),

- der Kontrolle des Sozialisationsauftrags durch die Sozialisationsagentur (systemfunktionalistische Ansätze) und als Variante davon mit der
- Kontrolle des Produkts 'akademische Arbeitskraft' durch den Hersteller 'Universität' (Kritische Theorie).

Die Funktionen einer Prüfung werden - mit unterschiedlicher Gewichtung - definiert als

- statuszuweisend,
- sozialisationsabschließend,
- studiensteuernd und
- sozial selektierend (bewertend).

Wenn wir die Prüfung als Problem bezeichnen, für das der Kandidat Problemlösungen finden muß, dann können wir die Statuszuweisung und die soziale Selektion als 'Zielzustand' betrachten.

Allgemein gehen wir davon aus, daß zur Bestimmung des Begriffs 'Problem' die objektiven Charakteristiken (und die Zusammenhänge zwischen ihnen) einer vorgegebenen Situation - z. B. der Prüfung - nicht ausreichen (vgl. KLIX, 1973). Notwendig ist, daß ein Zielzustand gegeben oder beschrieben sein muß, der aus der Anfangssituation abgeleitet oder erzeugt werden muß; sein Erreichen ist mit der Lösung des Problems identisch. Ferner gehört zu einem 'Problem', daß die Überführung der Ausgangssituation in den Zielzustand nicht automatisch gelingt, sondern eine Reihe von (beobachtbaren oder nicht-beobachtbaren) Operationen für diesen Zweck erst gesucht werden muß.

Je nach dem Ausmaß, in dem diese drei Bestimmungskomponenten eines Problems - Ausgangssituation, Zielzustand, verknüpfende Operationen - gegeben sind, können wir nach KLIX (1973) verschiedene Klassen von Problemstellungen unterscheiden:

- Ausgangssituation und die Menge der zulässigen und möglichen Operationen sind definiert, gesucht ist der Zielzustand (z. B. Schachspiel),
- Zielzustand und die Menge der zulässigen und möglichen Operationen sind definiert, gesucht ist die Ausgangssituation (z. B. einfache chemische Analysen) und
- Ausgangssituation und Zielzustand sind gegeben, gesucht sind die Operationen, die die Ausgangssituation in den Zielzustand überführen (z. B. Brückenkonstruktion).

Statuszuweisung, soziale Selektion und der Nachweis des Gebrauchs- und Tauscherts der eigenen Qualifikation sind der Zielzustand des problemlösenden Prüfungskandidaten.

Die Ausgangssituation ist für den problemlösenden Kandidaten mit der studiensteuernden Funktion von Prüfungen verknüpft. Der wichtigste Aspekt dabei läßt sich aus der Auffassung der Prüfung als Ritual ableiten: dem Prüfling wird jetzt (nicht vorher und nicht später) Gelegenheit gegeben, den Zielzustand unter den gegebenen Voraussetzungen zu erreichen. Der Prüfling muß also am Ende eines Studienabschnitts zu einem gegebenen Zeitpunkt und für einen gegebenen Zeitraum (Ausgangssituation) durch geeignete Operationen den Prüfer veranlassen, den Zielzustand der Statuszuweisung, der sozialen Selektion und des Qualifikationsnachweises in bestimmter Weise zu realisieren.

Somit besteht für den Kandidaten eine Problemstellung, in der Ausgangssituation und Zielzustand definiert und die erforderlichen Operationen zur Verknüpfung von beiden gesucht sind.

Als Bestimmungsstücke für die operative Verknüpfung haben wir gefunden (und nennen sie ohne Gewichtung):

- Überprüfung bestimmter Verhaltensmuster ('pattern variables'),
- Forderung eines bestimmten Rollenverhaltens ('role-making'),
- Forderung einer bestimmten (fach-)sprachlichen Diktion,
- die inferiore Position des Prüflings,
- die Prüflingsrolle des Ohnmächtig-Ausgelieferten,
- die Überprüfung homogener und mit denen der Prüfer übereinstimmender Wertorientierungen,
- die Internalisierung der quantifizierenden Bewertung durch Vorgesetzte,
- 'psyching out'.

Auffallend ist, daß drei Elemente einer Prüfung nur sehr selten aufgegriffen werden:

- niedrigere kognitive Lernziele sensu BLOOM et al. (1956, deutsch: 1974), also Prüfungsstoff im Sinne von '... der Student soll wissen, daß ...',
- die Person des Prüfers und
- Prüfungsmodi (schriftlich vs. mündlich).

Für die mangelnde Berücksichtigung einfacher kognitiver Lernziele sehen wir zwei potentielle Ursachen: Zum einen ist es möglich, daß der Abstraktionsgrad

der vorgestellten theoretischen Ansätze so hoch ist, daß einfache kognitive Lernziele nicht mehr konzeptualisiert werden, zum anderen kann es sein, daß für die Funktionen von mündlichen Prüfungen dieser Prüfungsstoff im engeren Sinne tatsächlich irrelevant ist. Beide Ursachen schließen sich nicht gegenseitig aus: Wenn wir über allgemeine Prüfungsfunktionen sprechen, dann müssen wir von Unterschieden im Detail absehen; detaillierte Unterschiede in diesem Sinne bestehen zwischen Prüfungen, für die Gebrauchs- und Tauschwert sensu KVALE gleich sind; also z. B. zwischen funktional äquivalenten Prüfungen des gleichen Berufsbildungsganges; für die Funktion von mündlichen Prüfungen sind einfache kognitive Lernziele in diesem Sinne irrelevant, zumal sie sich im Laufe der Jahre mit dem Fortschritt der Wissenschaft ändern. Relevant hingegen sind eher allgemeinere Lernziele wie Verhaltensmuster, Rollenverhalten, Sprachverhalten, Problemlösungsstrategien etc., die wir im Moment nur global angeben können und deren Beziehung zu unserem eigentlichen Anliegen - der Prüfungsangst - an dieser Stelle noch offen bleibt.

Die Person des Prüfers wird - wenn überhaupt - nur über den Rollenbegriff und hier im engeren Sinne über das Konzept des Sozialisationsagenten aufgegriffen. Während wir es bei schriftlichen Prüfungen - insbesondere bei solchen nach dem Multiple-Choice-System - für zulässig halten, von individuellen Prüfer-Eigenheiten abzusehen, ist dies bei mündlichen Prüfungen nicht automatisch plausibel. Die mangelnde Berücksichtigung der Person des Prüfers und die mangelnde Unterscheidung zwischen Prüfungs-

modi treffen hier zusammen.

Ein Prüfer ist jedoch auch Mitglied eines Berufsstandes (dem der Professoren), der sich selbst rekrutiert: ein 'Ruf' geht in der Regel von zukünftigen Kollegen aus. Ein Prüfer ist in seiner Berufslaufbahn daher verstärkt dem Prinzip der homogenen und übereinstimmenden Wertorientierung unterworfen, das sich gerade auf die prüfungsrelevanten Orientierungen bezieht. Von daher gesehen halten wir es für möglich, daß - bei allen sonstigen Unterschieden - sich Prüfer in der Prüfung gleichförmig verhalten, jedenfalls was die Funktionen und höheren Inhalte der Prüfungen betrifft.

Empirische Belege, die gegen unsere Annahme sprechen, konnten wir nicht finden. Auch in der Prüfungsangstliteratur wird von Prüferunterschieden bislang abgesehen, indem entweder Prüfungen bei einem einzigen Prüfer untersucht wurden (z. B. PIEHL, 1972) oder über alle Fächer und Prüfer Daten erhoben wurden (z. B. SCHEER & ZENZ, 1973). Wir wollen unsere Annahme der Prüfergleichheit deshalb heuristisch aufrechterhalten, gleichzeitig aber die Gelegenheit nutzen, erste Daten zu diesem Komplex zu sammeln (Hypothesengenerierungskomplex A).

Wir wollen zu unserer Globalhypothese zurückkehren, Prüfungsangst sei kein Resultat zufälliger individueller Lernprozesse. Wir haben Ansätze vorgestellt, nach denen Prüfungen Situationen sind, in denen der Kandidat unabhängig vom engeren Prüfungsstoff den Nachweis für das Erreichen bestimmter

übergeordneter Lernziele zu erbringen hat. Aus der Diskussion über unterschiedliche Systemfunktionalität von Prüfungen, die sich vom Verlauf des Studiums her unterscheiden lassen, hatten wir die Hypothese abgeleitet, daß Prüflinge in einer Abschlußprüfung im Schnitt bessere Zensuren erreichen als Prüflinge in einer Zwischenprüfung.

Aus der systemfunktionalistischen Unterscheidung verschiedener Prüfungsfunktionen folgt, daß Prüfungen für die Kandidaten auch einen unterschiedlichen Stellenwert haben. Bislang fehlt zu unserer Verblüffung ein empirischer Beleg, daß sich z. B. Kandidaten in Zwischen- und Abschlußprüfungen in den Motiven unterscheiden, die zur Prüfungsanmeldung führen. Wir wollen auch hier die Gelegenheit nutzen und hierzu erste Daten sammeln (Hypothesengenerierungskomplex B).

Wenn wir Prüfungen aus der Sicht des Kandidaten als Problemlösungsaufgabe konzipieren, dann nehmen wir an, daß Prüflinge Problemlösungsstrategien verwenden. Wir wollen sie Prüfungsstrategien nennen und die Diskussion im nächsten Kapitel aufnehmen.

Relativ undifferenziert bleibt vorerst unsere Beobachtung, daß die Vorstellungen von der Universität als totaler und unveränderbar vorgegebener Institution in einem empirischen Zusammenhang mit 'Prüfungsbelastungen' zu stehen scheint.

2.1.2 Die psychologischen Komponenten der mündlichen Prüfung

2.1.2.1 Die Interaktion in Prüfungen als Rahmen- bedingung für Prüfungsstrategien - symbolisch-kommunikative Betrachtungsebene

Wir haben ausgeführt, daß Prüfungen aus der Sicht des Kandidaten eine Problemsituation darstellen, in der es darum geht, Operationen zur Verknüpfung von Zielzustand und Ausgangssituation zu finden.

Zunächst und in erster Linie ist die mündliche Prüfung als eine Situation zu sehen, in der mindestens zwei Personen (Prüfer und Kandidat) miteinander kommunizieren. Wir wollen uns der analytischen Einfachheit wegen auf eine dyadische Prüfungskommunikation beschränken.

Wie bei jeder dyadischen Interaktion setzt sich die Kommunikationssequenz aus den Anteilen beider Partner zusammen. Die Kommunikation wird also nicht nur von einem Partner - dem Prüfer - gesteuert, sondern kann nur dann zustande kommen, wenn der andere Partner - der Kandidat - ebenfalls kommuniziert. Aus kommunikationstheoretischer Sicht steuert der Prüfling damit ebenfalls die Prüfung.

Der Part des Prüfers im System Prüfung beinhaltet zwei Funktionen: eine steuernde (er muß Fragen stellen, den Kandidaten ermuntern, unterbrechen usw.) und eine bewertende (er muß zum Schluß die Zensur festlegen).

Die bewertende Funktion impliziert die Aufgabe, aus der Interaktionssequenz Indizien für ein bestimmtes Notenniveau zu abstrahieren. Dies ist nur dann möglich, wenn eine Interaktion stattfindet, und um dies sicherzustellen, bedient sich der Prüfer der Steuerungsfunktion, d. h. er variiert das Niveau seiner Fragen so lange, bis er ein Niveau erreicht, für das vom Kandidaten ausreichend Indizien geliefert werden. In der Praxis sieht dies vermutlich so aus, daß zu Beginn Fragen mittlerer Schwierigkeit gestellt werden und in Abhängigkeit von den Antworten - den vom Kandidaten gelieferten Indizien - die weiteren Fragen ausgewählt werden.

Den Part des Prüfers können wir demnach - aus der Sicht des Prüfers - als Erkenntnisprozeß beschreiben, der einer anderen Klasse von Problemstellungen zuzuordnen ist: Ausgangssituation und die Menge der möglichen Operationen sind definiert, gesucht ist der Zielzustand.

Den Part des Kandidaten hatten wir bereits klassifiziert: Ausgangssituation und Zielzustand sind definiert, gesucht sind die verknüpfenden Operationen.

Oben hatten wir die Konzepte der Identifikation mit dem Prüfer, der Enkulturation und der Akkulturation dargestellt. Danach gleichen die Kandidaten mit fortschreitendem Studium ihr Verhalten mehr und mehr dem der Prüfer an. Nach KVALE (1972) ist eine positive Korrelation zwischen dem Ausmaß der

Angleichung und der Zensur zu erwarten.

Wir folgern daraus, daß die vom Kandidaten gesuchten Operationen (abstrakt) genau denen entsprechen, die der Prüfer (abstrakt) einsetzt. Der Kandidat muß also seinen Teil der Prüfungskommunikationssteuerung so auslegen, daß sie der des Prüfers gleicht. In kommunikationstheoretischer Sicht entspricht dies einer symmetrischen Interaktion: Prüfer und Kandidaten kommunizieren wie unter gleichberechtigten Kollegen. Genau in dem Umfang, in dem dem Kandidaten die Realisation einer symmetrischen Kommunikation gelingt, sollte er eine gute Zensur zu erwarten haben und sollten sich die beiden 'Zielzustände' angleichen.

Die vom Kandidaten gesuchten Operationen sind also dadurch gekennzeichnet, daß sie eine bestimmte Beziehung definieren: die symmetrische Interaktion.

Die Suche nach Beziehungskomponenten unterliegt Strategien. Wir meinen, daß sich Kandidaten darin unterscheiden.

Die Theorie der Gelernten Hilflosigkeit unterscheidet nun Individuen danach, welche Strategien sie gelernt haben. Die Unterscheidung geht aber einher mit Unterschieden in der emotionalen Reaktion.

Wir haben festgestellt, daß der Begriff Prüfungsangst neben dem Begriff der Prüfung auch die emotionale Reaktion der Angst als Bedeutungskomponente enthält.

Im folgenden Teil wollen wir darstellen, welche empirischen Ergebnisse zur Symptomatik der Prüfungsangst vorliegen, damit wir sie später mit den emotionalen Komponenten der Gelernten Hilflosigkeit vergleichen können.

2.1.2.2 Die Symptomatik der Prüfungsbelastung - organisch-physiologische, motorisch-ver- haltensmäßige und subjektive Betracht- ungsebenen

In einer Reihe von Studien wurde versucht, mittels faktorenanalytischer Techniken 'Prüfungsangstsyndrome' zu isolieren. Wir wollen uns der direkteren Übertragbarkeit wegen auf Studien aus dem deutschsprachigen Raum beschränken.

Die bekannteste der faktorenanalytischen Studien stammt von SCHEER & ZENZ (1973). Sie untersuchten 99 Studenten verschiedener Fächer vierzehn Tage vor und unmittelbar nach einer Zwischenprüfung mit einem ad-hoc konstruierten Prüfungsfragebogen, dem Gießen-Test (BECKMAN & RICHTER, 1972), je einem Fragebogen zum Studienverhalten und zur Universitäts-situation sowie einer Liste von Körpersymptomen. Eine Faktorenanalyse mit anschließender Varimax-rotation ergab sechs Faktoren:

1. Kontakt- und Selbstsicherheit (bzw. -unsicherheit),
2. Prüfungserleben (positiv bzw. negativ),
3. Sicherheit und Autonomie in der Arbeit (bzw. depressive Arbeitsstörung),
4. (kritische bzw. affirmative) Einstellung zur Uni-
versität,

5. Vegetative Reaktionsbereitschaft,
6. Leistungskonflikt.

HEINERTH (1972a) untersuchte Studentinnen der PH Bonn zwanzig Minuten vor einer mündlichen Prüfung mit einem selbstentwickelten Fragebogen. Eine Faktorenanalyse nach der Hauptachsenmethode mit anschließender Varimaxrotation ergab sechs Faktoren:

1. Depressivität,
2. Eßstörungen,
3. Aggressivität/Frustrabilität,
- 4, körperliche Beeinträchtigungen,
5. sozial-emotionelle Beeinträchtigung und
6. einen nicht interpretierten Faktor.

HECKHAUSEN (1977) fand bei einer Stichprobe von 70 Vordiplomanden der Psychologie folgende sechs varimaxrotierten Faktoren:

1. Inkompetenzerleben vs. Kompetenzerleben,
2. Mißerfolgserwartung,
3. negative vs. positive Selbstbewertung,
4. Aufgeregtheit vs. Gelassenheit,
5. Überforderungsgefühle vs. Spaß an Schwierigkeits-
meisterung,
6. distanziert-sachliche Reaktionskontrolle bzw.
mangelnde Reaktionskontrolle.

Während bei den ersten beiden Arbeiten ein theoretischer Hintergrund - soweit er sich auf die Auswahl der Items bezieht - nicht zu erkennen ist, geht HECKHAUSEN von der Aufmerksamkeitsinterferenztheorie WINES (1970) und der Zwei-Komponenten-These

der Ängstlichkeit aus (LIEBERT & MORRIS, 1967; MORRIS & LIEBERT, 1970; MORRIS & LIEBERT, 1969; DOCTOR & ALTMAN, 1969).

Aus der Gesamtschau der faktorenanalytischen Studien ergibt sich die Erkenntnis, daß Prüfungsangst mit Sicherheit kein unidimensionales Syndrom ist, sondern daß wenigstens Angst, Depression, unspezifische Körpersymptome und sozial-emotionelle Beeinträchtigungen zum Bild des Syndroms gehören.

PUCHTA & ZIOLKO (1969) beschrieben die Symptomatik von stationär behandelten Examenskandidaten:

"Die Symptomatik war in erster Linie durch die akut aufgetretenen Schwierigkeiten im Examen, insbesondere die plötzliche Angst mit völligem Leistungsveragen bestimmt. Zudem bestanden depressive Verstimmungen mit Weinkrämpfen bis zur Verzweiflungshaltung, völligem Antriebsmangel; in einem Fall war es zu einem Suizidversuch gekommen. Außer den skizzierten Beschwerden fanden sich diffuse vegetative Störungen." (PUCHTA & ZIOLKO, 1969, S. 312)

Weitere Variablen, die diese Gruppe kennzeichnen, sind Langzeitstudium, Druck von seiten der Familie, das Examen abzulegen, und Prüfungsängste in früheren, erfolgreich bestandenen Prüfungen.

Auch MALLESON (1957, 1959) beschreibt ein Angst-/Depressionsgemisch bei Examenskandidaten, die ein Student-Health-Center aufsuchen. Dominierende Merkmale sind Schlaflosigkeit, ständige Anspannung und Konzentrationsschwierigkeiten, Hoffnungslosigkeit und Gedanken, die kontinuierlich auf pessimistische Weise um das Examen kreisen.

ZIOLKO (1968, S. 61) nennt als häufiges Bild bei Examenskandidaten das der (depressiven) Erschöpfungsneurose mit charakteristischen Symptomen wie Apathie, 'Ich-Einschränkung', 'hochgradige Insuffizienz' usw.

Stützt man sich also auf klinische Beschreibungen Prüfungsängstlicher, so werden übereinstimmend Erschöpfung, Depressionen, Angst, Konzentrationsunfähigkeit mit Arbeitsstörungen, Apathie und Sorgen, die die Prüfung zum Inhalt haben, genannt.

Einen weiteren Ansatz, spezifische Probleme von Prüfungskandidaten zu erfassen, legen SCHEER & ZENZ (1973) vor. Sie interpretieren den ersten extrahierten Faktor der Hauptkomponentenlösung (also vor der Rotation) als 'Angstfaktor'. Danach hat der Kandidat (14 Tage vor der ersten Prüfung) vor der Prüfung Angst, hat es in Prüfungen schwer, läßt sich durch Aufregungen und Schwarzmalereien stören, ist in den Prüfungsfächern noch nicht sicher, hatte bereits vor dem Abitur Angst, erfährt schwerer persönliche Wertschätzung, fühlt sich in Prüfungen der Situation ausgeliefert, hat Angst, fertiggemacht zu werden, wird Pessimismus zur Schau stellen und sich vorsichtig zeigen.

In seinem Selbstbild fallen Züge der Ängstlichkeit, Bedrücktheit und sozialen Inkompetenz auf. Sein Arbeitsverhalten schildert der Prüfungsängstliche als unsicher, unstetig und zögernd. Vor der Prüfung empfindet er starke Körpersymptome, während der Prüfung ein Gefühl der Leere im Kopf und innere Unruhe. Nach der Prüfung bestätigt der Ängstliche ein hilfloses, pessimistisches, selbstkritisches und vorsichtiges Vorgehen.

Prüfungsängstliche definieren danach die Prüfung als eine soziale Situation, die über sie unkontrol-

lierbar hereinbricht, in der sie von einem übermächtigen Prüfer fertiggemacht werden, in der sie sich pessimistisch und selbstkritisch darstellen und sich dabei passiv und hilflos verhalten.

Ihre Interpretation des ersten extrahierten Faktors als 'Angstfaktor' legitimieren SCHEER & ZENZ (1973, S. 102) mit den hohen Ladungen der angstbezogenen Items. Methodenimmanent ist dabei, daß der erste extrahierte Faktor die meiste Varianz aufklärt, d. h. letztlich, daß auf diesem Faktor gerade die Variablen hoch laden, die am meisten zur Unterscheidbarkeit von Probanden beitragen. Ein Kandidat, der keine Prüfungsangst hat, wird sich also obiger Beschreibung entgegengesetzt darstellen.

Einen anderen Weg der Analyse von Prüfungsbelastungen beging TURKSON (1970), der 64 Studenten vor und nach einer Zwischenprüfung befragte. Er rechnete eine Faktorenanalyse vom Typ Q und extrahierte drei Faktoren, i. e. Typen von Kandidaten. Die erhaltenen Faktorladungen korrelierte er anschließend mit den Variablen und baute seine Interpretation der Typen auf der Annahme auf, daß eine gesicherte Korrelation zwischen dem Faktor (Typ) und der Variablen ein Indikator für typische Merkmale dieses Personentyps sei.

TURKSON macht keine Angaben, ob er mit einer rotierten oder unrotierten Matrix arbeitet.

Der erste Typ (ca. 24 % extrahierter Varianz) schildert sich danach als selbstsicher und nicht irri-

tierbar. Die Konzentrationsfähigkeit ist (nach eigenen Angaben) nicht gestört. Somatische Symptome treten kaum auf. Dieser Typ nennt folgendes Vorgehen als richtig, gibt an, so vorgehen zu wollen und so vorgegangen zu sein: Optimismus und Mut zeigen, selbstsicher auftreten, dem Prüfer ggf. widersprechen; dieser Typ gibt an, eher viel gesagt zu haben, wäre enttäuscht über eine Verzögerung des Termins und erlebte den Prüfer strenger als erwartet. Die Korrelation der Zensur mit den Faktorladungen dieses Typs ist nicht signifikant.

Bei diesem Typ ist selbstsicheres Verhalten geplant und realisiert worden. Seine Strategie ist mit wenigen Körpersymptomen assoziiert. Die Einschätzung der Prüfer als strenger als erwartet deutet an, daß die Prüflinge dieses Typs die Anforderungen durch den Prüfer, die Steuerung der Prüfung, leicht unterschätzten.

Der zweite Typ (ca. 13 % extrahierter Varianz) schildert sich als irritierbar, ablenkbar und sozial unsicher. Die von diesem Typ als richtig bewertete und eingehaltene Strategie ist durch die Variablen Pessimismus, Vorsicht und Selbstkritik sowie den Entschluß, auf keinen Fall zu widersprechen, auch wenn man recht zu haben glaubt, gekennzeichnet. Diese Gruppe erlebte während der Prüfung den Prüfer milder als erwartet und schnitt nach eigener Einschätzung besser als erwartet ab. Die Körpersymptome während der Prüfung sind mit einer Itemsausnahme (Mattigkeit vor der Prüfung) durchweg erhöht.

Bei diesem Typ vermuten wir eine passive Prüfungsstrategie und selbstunsicheres Verhalten. Die Erwartung der Nichtkontrollierbarkeit des Prüfungsverlaufs ist bei genauer Lektüre von TURKSONs Daten evident, ebenso wie die begleitenden vegetativen Symptome.

Für den dritten Typ nennt TURKSON (1970, S. 35) nur noch neun signifikante (von vermutlich insgesamt 109) Korrelationen. Dieser Typ ist u. a. durch eine vorsichtige Strategie mit dem Entschluß, eher wenig zu sagen, gekennzeichnet. Wegen der geringen Datenbasis bleibt dieser Typ von uns undiskutiert.

Auch TURKSONs Arbeit liefert reichlich Hinweise dafür, daß bestimmte Strategien mit bestimmten Belastungen einhergehen. Pessimismus, Vorsicht, Selbstkritik, Zurückhaltung und die Absicht, wenig zu sagen und dem Prüfer nicht zu widersprechen, scheinen mit emotionalen Belastungen, soweit sich diese in Körpersymptomen zeigen, assoziiert zu sein. Umgekehrt korreliert eine aktive, mutige, assertive Strategie mit geringen Belastungen.

Ein weiterer Ansatz, die Belastungen von Prüfungskandidaten zu erfassen, läuft über die Korrelationen mit dem Prüfungsergebnis, der Zensur.

TURKSON (1970) berichtet, daß schlechte Gesamtnoten mit folgenden Variablenausprägungen einhergehen:

In der Kindheit wurden Schulnoten und Zeugnisse zu wichtig genommen, das Verhalten eines Professors in der Prüfung ist nicht milder als sonst, man läßt sich bei der Arbeit nicht leichter ablenken als andere, man soll in der Prüfung eher wenig sagen und sich vorsichtig zeigen; nach der Prüfung geben Schlechtbewertete an, sie seien eher selbstkritisch aufgetreten und hätten schlechter als erwartet abgeschnitten.

SCHEER & ZENZ (1973) berichten, daß u. a. folgende Itemantworten mit gutem Abschneiden einhergehen:

Der gut Abschnaidende wäre enttäuscht über eine Verzögerung des Prüfungstermins, bereitet sich intensiver als andere vor, wird in den Prüfungen viel sagen, ist daran interessiert, andere zu übertreffen, legt es darauf an, andere zu lenken, bleibt leicht bei einer Sache, ist ehrgeiziger als andere, hat ein gutes Gedächtnis, meint, daß Studenten mehr Einfluß auf die Gestaltung der Universität bekommen müßten; während der Prüfung hatte er weniger Angst als erwartet, hat sich in der Prüfung der Situation gewachsen gefühlt, hat sich mutig gezeigt, erlebte die Prüfer milder als erwartet; vor der Prüfung sind die Symptome Leere im Kopf und Konzentrationsstörungen signifikant weniger ausgeprägt (11 andere vegetative Reaktionen korrelieren nicht mit der Zensur); während der Prüfung sind die Reaktionen niedergedrückte Stimmung, Leere im Kopf, innere Unruhe, Konzentrationsstörungen, Herzklopfen, Übelkeit und Mattigkeit signifikant weniger ausgeprägt.

Schlechte Zensuren gehen in beiden Arbeiten einher mit einer passiven, vorsichtigen Strategie und verstärkten Berichten über vegetative Reaktionen während der Prüfung.

Der von uns postulierte Zusammenhang zwischen Strategie und Zensur wird auch durch die Daten von HECKHAUSEN (1977) gestützt. HECKHAUSEN berichtet über Zusammenhänge zwischen dem Niveau der in Psychologie-Vordiplomprüfungen erreichten Zensur und den von ihm extrahierten Faktoren 'Inkompetenzerleben', 'Negative Selbstbewertung' und 'Mißerfolgserwartung'. HECKHAUSEN meint mit WINE (1971), daß der ängstliche Prüfungskandidat die ihm zur Verfügung stehende Zeit nicht voll für prüfungsrelevante Kognitionen nutzt, sondern einen Teil der Gedanken für Selbstzweifel verwendet. Je schlechter die erzielte Prüfungsnote - so die Ergebnisse von HECKHAUSEN - desto häufiger standen während der Prüfung ursachenanalytische Überlegungen

hinsichtlich internaler Faktoren (Unfähigkeit, Passivität) im Zentrum der berichteten Kognitionen, desto häufiger reflektierten die Kandidaten das Nichterreichen der gesetzten Norm, einen ungünstigen Prüfungsverlauf und Mißerfolgserwartungen. Letztere Variablen stehen in Zusammenhang mit externaler Kausalattribution der Zensur und den Gefühlen der Verärgerung und Unzufriedenheit nach der Prüfung.

2.1.2.3 Zusammenfassung und Diskussion

Zusammenfassend können wir sagen, daß die Emotion der 'Prüfungsangst' keine einheitliche Symptomatik hat. Das Stereotyp des zitternden und schwitzenden Kandidaten muß vielmehr ersetzt werden durch die Vorstellung eines Angst-/Depressionsgemisches, das Zittern, Erschöpfung, Passivität wie Überaktivität, Anspannung, Konzentrationsschwierigkeiten, Aggressivität wie Weinkrämpfe, Arbeitsstörungen wie soziale Kontaktstörungen umfaßt. Für den Klinischen Psychologen ergibt sich daraus die Konsequenz, bei 'Prüfungsängstlichen' sich weniger an der monosymptomatischen Tierphobie und mehr am 'panphobic syndrom' (MARKS, 1969) der Agoraphobie zu orientieren.

Der Verfasser zieht für sich zunächst die Konsequenz, das Wort 'Prüfungsangst' durch den Begriff der 'Prüfungsbelastung' zu ersetzen.

Soweit Analysen der 'Strategien' und Taktiken vorliegen, ergibt sich als Resultat, daß Prüfungsbe-

lastete sich der Situation Prüfung ausgeliefert fühlen, sie als unkontrollierbar erleben, das Gefühl haben, von einem Übermächtigen Prüfer fertig gemacht zu werden; dabei geben sie für vergangene wie zukünftige Prüfungen eine passive, vorsichtige, reaktive 'Strategie' an. Jemand ohne Prüfungsbelastungen schildert sich entgegengesetzt: Er hat das Gefühl, mitmischen zu können und erlebt den Prüfer zwar einflußreicher als erwartet, aber nicht als Übermächtigen Sadisten; dabei gibt er für zukünftige wie vergangene Prüfungen eine aktive und mutige 'Strategie' an.

Interessanterweise wird eine aktive und mutige 'Strategie' in Beziehung zu guten Zensuren gesetzt, eine passive und vorsichtige wird mit schlechten Zensuren assoziiert; gleichzeitig finden wir empirische Hinweise, daß aktive und mutige Strategien mit geringen Prüfungsbelastungen, passive und vorsichtige mit stärkeren Belastungen einhergehen.

Bei der Analyse des Problems Prüfung aus der Sicht des Kandidaten haben wir hervorgehoben, daß es für den Kandidaten in erster Linie um Operationen innerhalb des Kommunikationssystems Prüfung geht, die eine symmetrische Beziehung herstellen. Nur dann kann er seinen Zielzustand - eine möglichst gute Zensur - erreichen. Diese symmetrische Beziehung taucht unseres Erachtens in den Studien zur Symptomatik immer wieder unter den Items 'sich mutig, aktiv, widersprechend verhalten, mitmischen u. a.' auf - ein Gleicher spricht mit einem Gleichen.

Das Gegenteil der symmetrischen Beziehung - die

komplementäre (vgl. WATZLAWICK, BEAVIN & JACKSON, 1969) - können wir auf Items wie 'sich ausgeliefert fühlen, fertiggemacht werden, vorsichtig und passiv vorgehen, wenig sagen u. a.' beziehen. Dabei finden wir eine empirische Assoziation der komplementären Interaktion sowohl mit Prüfungsbelastungen wie mit schlechten Zensuren.

Dies führt uns zurück zu den sozialwissenschaftlichen Funktionen von Prüfungen. Dort hatten wir geschrieben, daß die Sicht der Universität als totaler und unveränderbar vorgegebener Institution in einem empirischen Zusammenhang zu Prüfungsbelastungen stehe. Die hier zitierten Korrelationsstudien ergeben übereinstimmend, daß der Glaube an die Unbeeinflussbarkeit von Prüfungen mit einem Angst-/Depressionsgemisch einhergeht. Wenn wir diese beiden Punkte vereinigen und berücksichtigen, daß der Kandidat gerade in einer Prüfung mit der Organisationsstruktur der Institution Universität konfrontiert wird und sich mit ihr auseinanderzusetzen hat, dann kommen wir zur Schlußfolgerung, daß nicht nur die objektive - offene oder geschlossene - Organisationsstruktur, sondern auch das subjektive Erleben einer Institution als geschlossen oder offen in Beziehung zu Prüfungsbelastungen, Strategien und Leistungen stehen muß.

Wir nehmen an, daß sich Kandidaten in dem Ausmaß unterscheiden, in dem sie Prüfungen als beeinflussbar erleben. 'Selbstsichere' realisieren dann eine symmetrische Interaktion, 'Unsichere' eine komplementäre; 'Selbstsichere' erreichen gute Zensuren,

'Unsichere' schlechte; 'Selbstsichere' leiden kaum unter Prüfungsbelastungen, 'Unsichere' sehr viel stärker.

Offen bleibt an dieser Stelle freilich, welche psychologischen Mechanismen den bis hier beschriebenen empirischen Zusammenhängen zugrunde liegen. Wir konstatieren lediglich eine empirische Dreiecksbeziehung zwischen Strategien, Belastungen und Prüfungsleistungen.

Wir können auch keine Aussagen zu Unterschieden zwischen Kandidaten in Prüfungen mit unterschiedlicher Funktion ableiten, die über die bereits formulierten Hypothesen hinausgehen.

Für die Prävention der Prüfungsbelastung ergibt sich bis jetzt ein Gesichtspunkt, der je nach der Seite des Tisches für wichtig (Kandidaten) oder trivial (Prüfer) gehalten werden kann: In Prüfungen geht es offenbar weniger um das, was ein Kandidat weiß, sondern allein um das, was er sagt. Für einen Prüfer ist es trivial, daß die Zensur ausschließlich mit den Äußerungen des Kandidaten begründet wird; für einen Kandidaten ist die Feststellung wichtig, daß eine Zensur nicht auf das bezogen werden kann, was er gesagt haben könnte. Weitergehende Erkenntnisse über die Prävention der Prüfungsbelastung können wir bislang nicht sehen.

Wir meinen nun, daß die Theorie der Gelernten Hilflosigkeit uns hier nicht nur weiterhilft, sondern auch eine theoretische Basis für die bislang dargestellten komplexen Beziehungen liefert.

Da von ihr noch keine kritische - im deutschsprachigen Raum noch gar keine - Darstellung vorliegt, wollen wir auf sie ausführlicher eingehen, als vom Thema her naheliegt. Der eilige Leser möge dies dem Verfasser verzeihen und sich auf die Zusammenfassungen am Schluß jeden Abschnitts beschränken.

2.2 Die Theorie der Gelernten Hilflosigkeit - Generalisierung und Differenzierung der Theorie in drei Phasen

Die Theorie der Gelernten Hilflosigkeit ist eine kognitive Theorie, die die Terminologie der behavioristischen Lerntheorien verwendet. Sie erhebt den Anspruch, ein Modell für die Ätiologie von depressiven Reaktionen zu liefern.

Das Konzept der Hilflosigkeit - und die dazugehörigen Gegenkonzepte wie Kontrolle, Kompetenz, Effektivität u. a. - waren schon vor SELIGMANs erster Veröffentlichung zu diesem Thema (SELIGMAN & MAIER, 1967) in der Psychologie bekannt und wurden auch schon vor ihm (SELIGMAN, 1973) auf emotionales Geschehen übertragen.

Bereits ADLER (1958) postuliert eine allgemeine Beziehung zwischen Macht (= Kontrolle) und neurotischem Verhalten. WHITES (1959) Konstrukte der Kompetenz und Effektivität ähneln ADLERs Machtstreben. Hilflosigkeit als Folge fehlender Kontrolle führt zu Schizophrenie (BATESON, JACKSON, HELEY & WEAKLAND, 1956, deutsch: 1970) und Depression (SELIGMAN, 1974, 1975). Definieren wir Hilflosigkeit als einen Zustand, in dem ein Individuum keine geeigneten

(Coping-) Reaktionen zur Verfügung hat, dann können wir nach LAZARUS (1966) und MANDLER (1972) Hilflosigkeit mit Angst gleichsetzen. Nach der Theorie von MANDLER & WATSON (1966) kann Hilflosigkeit zu den verschiedensten Emotionen führen.

Hilflosigkeit (als Zustand fehlender Kontrolle) wurde ferner in Beziehung gesetzt zu Magengeschwüren (WEISS, 1968), physischen Krankheiten (ENGEL, 1968; SCHMALE, 1971), plötzlichem Tod (RICHTER, 1959; GREENE, GOLDSTEIN & MOSS, 1972), zu sozialen Klassenunterschieden (JOE, 1971, S. 623 f.) und zu Terrorismus (ZIMBARDO, 1969).

SELIGMANs Verdienst war es zunächst, Hilflosigkeit lerntheoretisch so zu definieren, daß sie unter Laboratoriumsbedingungen untersucht werden konnte. Damit war ein Zugang mit der z. Z. vorherrschenden Forschungsmethode über einen z. Z. vorherrschenden Erklärungsansatz geschaffen.

Dann scheute er sich nicht, den so gewonnenen Ansatz spekulativ und spektakulär auf einige bislang lerntheoretisch und experimentell vernachlässigte Gebiete zu übertragen. Die größte Resonanz fand dabei seine Analyse depressiven Verhaltens (beginnend mit SELIGMAN, 1973).

Parallel wurde die LH-Theorie weiterentwickelt, wobei wir bis heute deutlich drei Phasen (in dieser Arbeit Phase I, Phase II und Phase III genannt) unterscheiden können.

Gemeinsames Element dieser drei Phasen ist eine ebenso simple wie plausible kognitive Lerntheorie, wonach der Organismus Informationen über die Auswirkungen seines Verhaltens (= Kontingen-

zen) im Gedächtnis speichert und das Gespeicherte irgendwann später sein Verhalten beeinflusst.

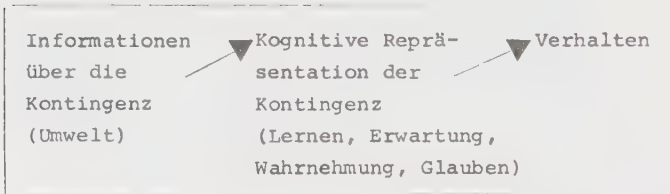


Abbildung 1: Die Dreikomponententheorie von
SELIGMAN (übers. aus SELIGMAN,
1975, S. 17)

"A man or an animal must begin with information about the contingency of outcome upon response. This information is a property of the organism's environment, not a property of the receiver... The information about the contingency must be processed and transformed into a cognitive representation of the contingency. Such a representation has been variously called 'learning', 'perceiving' or 'believing'... I prefer to call the representation the expectation..." (SELIGMAN, 1975, S. 47f.; vgl. MAIER & SELIGMAN, 1976, S. 17f.)

Der Lernprozeß beginnt mit Informationen über Kontingenzen in der Umwelt. Diese werden gespeichert; wir wollen SELIGMANs Terminologie übernehmen und das Resultat der Speicherung künftig 'kognitive Repräsentation' nennen. Die kognitive Repräsentation beeinflusst schließlich das Verhalten.

Übertragen auf den Problemlösungsprozeß mit seinen drei Bestimmungskomponenten Ausgangssituation, Zielzustand und Menge der möglichen und zulässigen Ope-

rationen bedeutet dies, daß der Organismus entweder die notwendigen Informationen bereits kognitiv repräsentiert hat (und nur noch ihre Struktur oder Sequenz herausfinden muß) oder daß er durch aktives Eingreifen in die Umwelt sich die Informationen über die Kontingenzverhältnisse erst verschaffen muß.

In den drei Phasen der Entwicklung der LH-Theorie werden diese Möglichkeiten unterschiedlich gewichtet.

2.2.1 Phase I: Beginn mit Tierexperimenten

2.2.1.1 Allgemeine Darstellung der Theorie

Zentrales Konzept der LH-Theorie ist die Kontrollierbarkeit (bzw. Unkontrollierbarkeit) von Verstärkern.

Verstärker werden von SELIGMAN (1975) und MAIER & SELIGMAN (1976) als Ergebnis einer freiwilligen Reaktion ('voluntary response') definiert. Eine freiwillige Reaktion ist dadurch definiert, daß sie durch Belohnung oder Bestrafung modifiziert werden kann; sie wird abgegrenzt gegen Reflexe, Instinkte und Tropismen.

Die Modifikation einer freiwilligen Reaktion nennt sich operantes Konditionieren. Verfügt ein Organismus während des operanten Konditionierens über eine Reaktion, die zu einem Ergebnis - dem Ver-

stärker - führt, dann ist dieses Ergebnis kontrollierbar. Wenn ein Organismus keine Reaktion realisieren kann, die zu einem Ergebnis führt, dann ist dieses Ergebnis unkontrollierbar.

Die Unkontrollierbarkeit eines Ergebnisses definiert Hilflosigkeit.

Wenn es - wie z. B. beim Klassischen Konditionieren - kein Ergebnis gibt, das kontrollierbar wäre, dann ist dies auch Hilflosigkeit.

"What defines a Pavlovian experiment and distinguishes it from an operant experiment is precisely helplessness. No response, conditioned or otherwise, is allowed to change the CS or US in classical conditioning; whereas in an operant study, there must be some response that obtains reward or alleviates punishment. Put another way, in instrumental learning the subject has a voluntary response that controls outcomes, but in Pavlovian conditioning he is helpless." (SELIGMAN, 1975, S. 12f.)

Die beim operanten Konditionieren ablaufenden Prozesse werden nach SELIGMAN (1975) durch den Reaktionskontingenzraum erfaßt (s. Abbildung 2).

Auf der x-Achse wird die Wahrscheinlichkeit K abgetragen, mit der irgendein Ergebnis C_j auf irgendeine Reaktion R_i erfolgt. (Abweichend von den Veröffentlichungen SELIGMANs orientieren wir uns an der Notation von KANFER & SASLOW, 1965, 1969). Kontinuierliche Verstärkung entspricht demnach $K = 1.0$; ist K kleiner als 1.0 (z. B. $K = 0.7$) so sprechen wir von partieller oder diskontinuierlicher Verstärkung.

Auf der Ordinate ist die Wahrscheinlichkeit K abgetragen, mit der irgendein Ergebnis C_j folgt, wenn R_i nicht auftritt ($= \bar{R}_i$). Hier sprechen wir von differentielltem Verstärken.

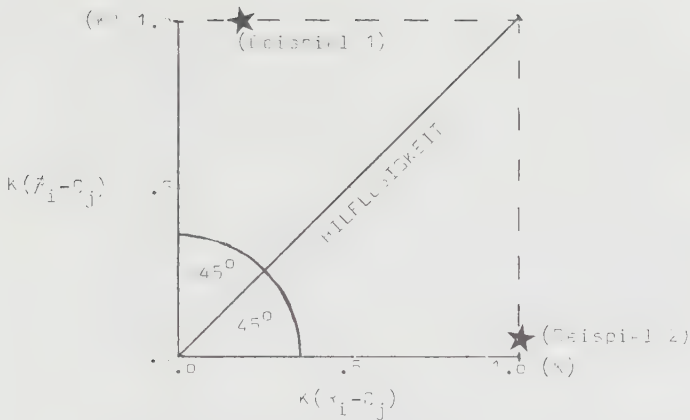


Abbildung 2: Der Reaktionskontingenzzraum
(Erläuterungen im Text)

Ein Beispiel für differentielles Verstärken wäre, wenn eine Mutter ihr Kind mit einem Eis dann belohnt, wenn es beim Essen nicht gekleckert hat. Wir wollen annehmen, daß die fiktive Mutter dies immer tut, sofern die o. a. Bedingungen erfüllt sind ($K = 1.0$). Natürlich wird es dann vorkommen, daß das Kind ein Eis erhält, obwohl es das Tisch Tuch beschmutzte, nämlich dann, wenn die Mutter das Kleckern nicht sieht (Beispiel 1).

Ein Beispiel für kontinuierliches Verstärken wäre, wenn eine Ehefrau ihrem Ehemann immer dann einen Kuß gibt, wenn er ihr Blumen mitbringt ($K = 1.0$). Wahrscheinlich wird es vorkommen, daß der Ehemann einen Kuß erhält, obwohl die Blumen nicht für seine Frau, sondern für einen geplanten Besuch gedacht waren, nämlich dann, wenn die Frau glaubte, die Blumen seien für sie bestimmt (Beispiel 2).

In der LH-Theorie wird nun angenommen, daß der Reaktionskontingenzzraum auch der kognitiven Repräsentation des Zusammenhangs zwischen Reaktion und Er-

gebnis entspricht. Der Reaktionskontingenzzraum bietet demnach nicht nur ein Modell der objektiven Kontingenzzverhältnisse in der Umwelt (vgl. die beiden o. a. Beispiele), sondern ist auch ein Modell für die kognitive Repräsentation der Kontingenzen. "We assume that organisms are sensitive to variations along both dimensions conjointly, and the empirical meaning of this assumption is that systematic changes in behavior should occur with systematic changes along both dimensions." (MAIER & SELIGMAN, 1976, S. 5).

Diese Annahme ist Voraussetzung für die genaue Definition von Kontrollierbarkeit und Hilflosigkeit.

"Consider the 45° line of a response contingency space. At any point on this line, the probability of an outcome is the same whether or not a given response occurs, the outcome is independent of that response. When this is true of all voluntary responses, the outcome is uncontrollable. Conversely, if the probability of an outcome when some response occurs is different from the probability of the outcome when that response doesn't occur, then that outcome is dependent on that response: the outcome is controllable... a rigorous definition under which helplessness occurs: a person or an animal is helpless with respect to some outcome when the outcome occurs independently of all his voluntary responses." (SELIGMAN, 1975, S. 16f.)

Wenn wir Bezug nehmen auf Abbildung 2, dann ist Hilflosigkeit definiert durch

$$K(R_i - C_j) = K(R_i - C_j) ; i = 1, 00 \text{ (Hilflosigkeit).}$$

Kontrolle, d. h. fehlende Hilflosigkeit oder auch

nach SELIGMAN (1975, S. 55) Kompetenz, ist definiert durch

$$K(R_i - C_j) \neq K(\bar{R}_i - C_j) ; i = 1,00 \text{ (Kontrolle, Kompetenz).}$$

Mit anderen Worten: Ein Individuum ist dann nicht hilflos, ist dann kompetent, wenn es ein bestimmtes Ergebnis durch irgendeine Reaktion herbeiführen oder verhindern kann (vgl. MAIER & SELIGMAN, 1976, S. 5f.)

Ein Organismus wird nun manchmal in eine Situation geraten, in der die Ergebnisse (Verstärker) abhängig sind von seinem Verhalten. In der LH-Theorie wird postuliert, daß der Organismus diese Kontingenz lernt. Wir nennen diesen Lernprozeß 'operantes Konditionieren'.

Ein Organismus wird sich aber auch in Situationen wiederfinden, in denen die Ergebnisse unabhängig sind von seinem Verhalten. In der LH-Theorie wird postuliert, daß der Organismus auch diese Kontingenz lernt.

Der Lernvorgang beginnt mit Informationen über Kontingenzen, die als Situationsdeterminanten, als Eigenschaften der Umwelt, als Set von Stimuli, jedenfalls nicht als Leistung des Individuums (MAIER & SELIGMAN, 1976, S. 17) zu betrachten sind.

Diese Informationen werden gespeichert. Der Organismus lernt (= entwickelt die Erwartung), daß Verhalten und Ergebnis unabhängig oder abhängig voneinander sind. SELIGMAN (1975) und MAIER & SELIGMAN (1976)

lassen es offen, wie es genau zur kognitiven Repräsentation kommt und welche Prozesse im einzelnen involviert sind. Zu diesem Punkt verweisen sie auf KELLEY (1967), WEINER et al. (1971), LAZARUS (1966) und andere. Wichtig ist allein (in der LH-Theorie), daß es zur kognitiven Repräsentation kommt.

Der von SELIGMAN und seinen Koautoren erforschte Lernvorgang beginnt mit einer Situation, in der der Organismus hilflos ist. Die Bedingungen für diese Hilflosigkeit, die Unkontrollierbarkeit der Ergebnisse, werden kognitiv repräsentiert. Die Autoren sprechen deshalb von Gelernter Hilflosigkeit (LH).

2.2.1.1.1 Die drei Aspekte der Gelernten Hilflosigkeit: kognitives, motivationales und emotionales Defizit

LH als Konstrukt ist nicht direkt beobachtbar, sondern muß über die Auswirkungen auf das Verhalten erschlossen werden.

"In general, when an organism experiences uncontrollable events, three deficits often ensue: motivational, cognitive, and emotional. a) The motivation to respond in the face of later aversive events seems to wane. b) Moreover, even if the subject does respond and the response succeeds in producing relief, the subject often has difficulty learning that the response worked. c) Finally, emotional balance may be disturbed; depression and anxiety, measured in a variety of ways, may predominate." (MAIER & SELIGMAN, 1976, S. 7)

Der LH-Effekt hat demnach drei Aspekte:

1. ein motivationales Defizit; dieses wird durch eine sinkende Frequenz operanten Verhaltens operationalisiert. Es deutet sich an, daß die Autoren unter verringerter Motivation ein verringertes Ausmaß mobilisierter Energie oder Aktivität in einem unspezifischen Sinne verstehen;
2. ein kognitives Defizit; LH unterminiert die Fähigkeit des Organismus, einen Zusammenhang zwischen Reaktion und Ergebnis zu lernen. Ein extrem stark ausgeprägter LH-Effekt sollte diese Fähigkeit dann auf Null reduzieren, selbst wenn die Ergebnisse kontrollierbar sind;
3. ein emotionales Defizit; als Folge der Konfrontation mit unkontrollierbaren Ergebnissen können Angst und Depression, aber auch Magengeschwüre (WEISS, 1968), Verlust von Appetit (MOWRER & VIEK, 1948), Urinieren und Defäkieren bei Hunden (SELIGMAN, 1975) auftreten.

Obiges Zitat enthält eine Einschränkung, die den bislang zitierten theoretischen Formulierungen nicht entspricht: die Folgen des Lernprozesses - wir wollen sie von jetzt an mit 'LH-Effekt' bezeichnen - werden auf aversive Ergebnisse beschränkt.

In dem bislang Dargestellten hielten wir uns relativ eng an die üblichen Darstellungen der LH-Theorie (SELIGMAN, 1975, MAIER & SELIGMAN, 1976, u. a.). In verkürzter Form lauten sie: Unter Bedingungen objektiver Unabhängigkeit von Reaktion und Verstär-

ker kommt es zur kognitiven Repräsentation dieser Unabhängigkeit; als Folge davon sind emotionale, kognitive und motivationale Defizite zu beobachten!

Da in der LH-Theorie explizite Aussagen über die bei der kognitiven Repräsentation involvierten Prozesse verweigert werden, sondern es allein genügt, daß es zur kognitiven Repräsentation gekommen ist (man erkennt dies an den Defiziten), gelangen wir zu der Schlußfolgerung, daß in die LH-Theorie möglicherweise tautologische Schlüsse eingebaut sind: Von den objektiven Bedingungen wird auf kognitive Zwischenprozesse der 'Gelernten Hilflosigkeit' geschlossen, die den LH-Effekt zur Folge haben; die kognitiven Zwischenprozesse, die den LH-Effekt zur Folge haben, kann man nun aus der Tatsache erschließen, daß in ihrer Folge der LH-Effekt auftritt.

Dies würde bedeuten, daß wir auf die Kognitive Repräsentation als Element der Theorie verzichten können und die objektiven Bedingungen direkt auf das beobachtbare Verhalten beziehen können.

Dies ist lediglich in Phase I möglich!

In allen uns bekannten Überblicksarbeiten wird der Reaktionskontingenzraum dargestellt und zum LH-Effekt so in Beziehung gesetzt, daß ein flüchtiger Leser den Eindruck haben kann, als seien die kognitiven Zwischenprozesse lediglich der Vollständigkeit halber aufgeführt. Diese Darstellungen können leicht irreführen: zentrales Element der

LH-Theorie ist die kognitive Repräsentation der Unabhängigkeit von Reaktionen und Verstärkern!

Wir wollen diesen Punkt an dieser Stelle (noch) nicht weiter diskutieren.

2.2.1.1.2 Das Immunisierungsparadigma

SELIGMAN & MAIER (1967) zeigten als erste Bedingungen auf, unter denen der LH-Effekt nicht auftritt. Sie gingen von der Annahme aus, daß der dem LH-Effekt zugrunde liegende Mechanismus auch umkehrbar sein müsse: Vorerfahrungen mit unkontrollierbaren Ergebnissen behindern das Lernen der Abhängigkeit von Reaktion und Ergebnis, also müßten Vorerfahrungen mit kontrollierbaren Ergebnissen das Lernen der Unabhängigkeit von Reaktion und Ergebnis behindern (mit ihm interferieren). Sie wiesen dies in einem Experiment mit Hunden nach. SELIGMAN, MARQUES & RADFORD (zit. nach SELIGMAN, 1974) replizierten diese Ergebnisse. SELIGMAN (1975) subsumiert sie unter dem Begriff der 'Immunisierung' und setzt sie in Beziehung zu prophylaktischen Maßnahmen.

Für die experimentelle Untersuchung folgt daraus, daß ein LH-Effekt seltener bei Tieren auftreten sollte, die Vorerfahrungen mit kontrollierbaren Ergebnissen haben, wenn man sie mit Tieren vergleicht, die geringe oder keine Gelegenheit hatten, Kontrollerwartungen auszubilden. SELIGMAN & GROVES (1970) verglichen Hunde, die einzeln und bestens

versorgt im Käfig aufwachsen, mit Hunden unbekannter Vorgeschichte (= erhöhte Wahrscheinlichkeit der Kontrollerwartung) und fanden dies bestätigt. Analoge Ergebnisse berichten für Hunde LESSAC & SOLOMON (1969) und für Ratten SELIGMAN, ROSSELINI & KOZAK (1975; vgl. dazu SELIGMAN, 1975, S. 134f.)

2.2.1.1.3 Die Unterscheidung von drei methodischen Ansätzen

Diese Ergebnisse zeigen, daß bestimmte Vorerfahrungen das Verhalten und Erleben in zeitlich späteren Situationen auf bestimmte Weise determinieren: Vorerfahrungen mit kontrollierbaren Ergebnissen beugen dem Auftreten des LH-Effekts vor, Vorerfahrungen mit unkontrollierbaren Ergebnissen begünstigen sein Auftreten.

Das zentrale Problem der LH-Forschung ist die Operationalisierung der Vorerfahrung.

Die erste Möglichkeit ist, die Versuchspersonen oder Versuchstiere nach der Wahrscheinlichkeit ihrer Vorerfahrung zu klassifizieren. Wie wir oben gesehen haben, zeigt sich der LH-Effekt bei Hunden und Ratten, deren Vorgeschichte bekanntermaßen wenig Gelegenheit bot, Ergebnisse zu kontrollieren, stärker als bei Tieren, deren Vorgeschichte unbekannt war. Von ersteren war bekannt, daß sie noch nicht einmal Gelegenheit hatten, sich die Nahrung selbst zu suchen; sie wurden in ihrem Käfig bestens, i. e. regelmäßig und reichlich, versorgt. Dies mag auch für einen Teil der anderen Tiere gegolten haben, aber

wahrscheinlich gab es unter ihnen eine größere Gruppe, die halbwild oder wild aufgewachsen war.

Allgemein können wir folgendes annehmen: Wenn die Wahrscheinlichkeit kontrollierender Vorerfahrung steigt, dann sinkt die Wahrscheinlichkeit für das Auftreten eines LH-Effekts; wenn die Wahrscheinlichkeit kontrollierender Vorerfahrung sinkt, dann steigt die Wahrscheinlichkeit für das Auftreten eines LH-Effekts. Wenn wir Versuchspersonen mit unterschiedlicher Vorerfahrung klassifizieren, dann können wir unterschiedliche Wahrscheinlichkeiten für das Auftreten eines LH-Effektes in den Klassen vorhersagen.

Eine zweite Möglichkeit zur Erfassung der Vorerfahrung ist, die aus ihr resultierenden Erwartungen zu messen. Diese Möglichkeit beinhaltet selbst zwei Varianten:

- a) Wir definieren und messen einen LH-Effekt und schließen retrospektiv auf entsprechende Vorerfahrungen und
- b) wir gehen den Weg der Konstruktvalidierung, d. h. wir formulieren über Metatheorien Hypothesen, die sich nicht direkt aus der LH-Theorie ableiten lassen.

Für die zweite Variante bieten sich im Humanexperiment z. B. das Konzept des Locus of Control (JOE, 1971; LEFCOURT, 1966; ROTTER, 1966; eine komplette Bibliographie legen THROOP & MacDONALD, 1971, vor), das Konzept des Motivkonflikts (FISCH, 1969, 1970, 1971; PIEHL, 1972; EPSTEIN & FENZ, 1962, 1965; FENZ & EPSTEIN, 1968, für eine allgemeinere Diskus-

sion vgl. HECKHAUSEN, 1974, 1975) und das Konzept der Kausalattribution (WEINER et al., 1971; DWECK, 1975; NISBETT & SCHACHTER, 1966; ROSS, RODIN & ZIMBARDO, 1969; STORMS & NISBETT, 1970; VALINS & NISBETT, 1971; PIEHL, 1976; WEINER & KUKLA, 1970) an.

Zunächst jedoch wollen wir bei den Forschungen SELIGMANS und seiner Mitstreiter bleiben und die dritte Möglichkeit der Operationalisierung diskutieren: das Training von 'Vorerfahrung'.

2.2.1.1.4 Hauptforschungsmethode: Die Realisation unterschiedlicher Vorerfahrungen im triadischen Design

Neben der logischen Wahrscheinlichkeitsklassifikation und der direkten oder indirekten Messung der Vorerfahrung gibt es eine - aus der Sicht der Grundlagenwissenschaftler sicherlich die überzeugendste - weitere Möglichkeit: die Vorerfahrung wird im Experiment realisiert. Diese wird von LH-Forschern bevorzugt.

Sie verwenden dafür das 'triadische Design'.

Im triadischen Design finden zwei Versuchsdurchgänge statt, zwischen denen eine für die Versuchstiere oder Versuchspersonen mehr oder weniger deutliche Unterbrechung liegt.

In der Pretestphase werden die unterschiedlichen Vorerfahrungen realisiert, d. h. die Kontingenzen von Ergebnissen werden kontrollierbar (K-Gruppe),

Bezeichnung der Versuchsgruppe	Pretestphase	Unterbrechung	Testphase
Controllable Group (= 'K-Gruppe')	Ergebnis ist kontrollierbar		Ergebnis ist kontrollierbar
Uncontrollable Group (= 'UK-Gruppe')	Ergebnis ist unkontrollierbar		Ergebnis ist kontrollierbar
Kontrollgruppe ('naiv', daher 'N-Gruppe')	Versuchsdurchgang entfällt, oder: Pseudotreatment		Ergebnis ist kontrollierbar

Abbildung 3: Das triadische Design der LH-Forschung

unkontrollierbar (UK-Gruppe) oder nicht realisiert (naive Individuen, N-Gruppe). Beispiele werden wir weiter unten darstellen. Die Auswirkungen der Pretestphase (vereinzelt auch Trainingsphase genannt) werden nach einer Unterbrechung in der Testphase geprüft, wobei dann alle Gruppen den gleichen Versuchsbedingungen unterworfen werden.

Der Vergleich der Testphasenergebnisse ist Grundlage für die Beurteilung eines oder mehrerer Aspekte des LH-Effektes. Da die Realisation der Versuchsbedingungen in der Testphase für alle drei Gruppen gleich ist, können Gruppenunterschiede auf Unterschiede der Treatments in der Pretestphase zurückgeführt werden.

In der K-Gruppe erhalten die Versuchstiere oder Versuchspersonen Aufgaben, bei denen Reaktion und Ergebnis (Verstärker) abhängig sind. In der UK-Gruppe werden die Versuchstiere oder Versuchspersonen mit Bedingungen konfrontiert, in denen Reaktion und Ergebnis unkontrollierbar sind. Bei Tierexperimenten wird jedes Versuchstier der UK- mit einem Versuchstier der K-Gruppe gekettet ('yoked') und erhält die Verstärker in der gleichen Stärke. Zeitpunkt und Kontingenz richten sich jedoch nach dem Verhalten der Versuchstiere der K-Gruppe, während die Versuchstiere der UK-Gruppe die Verstärker nicht beeinflussen können. Unterschiede zwischen diesen beiden Gruppen können demnach nicht auf Adaptation, Response Cost, Stressorintensität o. ä. zurückgeführt werden. Die Gruppen unterscheiden sich nach SELIGMAN und Mitarbeitern lediglich durch die Kontrollierbarkeit der Ergebnisse.

Die N-Gruppe wird keinem oder einem irrelevanten Treatment unterworfen. Diese Versuchstiere oder Versuchspersonen können keinerlei Erfahrungen mit Kontingenzen sammeln.

Die LH-Theorie sagt nun daraus, daß in der UK-Gruppe sich einer oder mehrere Aspekte des LH-Effektes zeigen und sich die UK-Gruppe von der K-Gruppe und der N-Gruppe signifikant unterscheidet.

In der Theorie des triadischen Designs (MAIER & SELIGMAN, 1976; SELIGMAN, 1975; SELIGMAN et al., 1971) werden keine Vorhersagen über Unterschiede zwischen N- und K-Gruppe gemacht.

Dabei wäre es möglich, daß in Analogie zur Theorie des Sicherheitssignals (SELIGMAN & MEYER, 1970; SELIGMAN, MAIER & SOLOMON, 1971; SELIGMAN, 1968, zusammenfassend SELIGMAN, 1975, S. 107f.) sich nicht nur ein learning set der Hilflosigkeit (SELIGMAN, 1975, S. 154), sondern auch ein learning set der Kompetenz ausbildet, daß Vorerfahrungen mit kontrollierbaren Ereignissen die Bewährung in neuen Situationen erleichtert. BLACK (1977) geht sogar so weit zu fragen, ob nicht unter bestimmten Umständen 'learned omnipotence' auftreten kann. Letzteres würde bedeuten, daß sich UK- und N-Gruppe voneinander, beide aber signifikant von der K-Gruppe unterscheiden.

CAMPBELL & STANLEY (1963, 1970) unterscheiden in ihrem Grundsatzreferat die 'Kontrollgruppe' von der 'Versuchsgruppe' durch das fehlende Treatment. In der mathematischen Formalisierung entspricht dies der Aufspaltung der empirischen Werte für die 'Kontrollgruppe' in 'wahren Wert' und einen Betrag der 'Fehlervarianz', der Werte der 'Versuchsgruppe' in 'wahren Wert', 'Fehlervarianz' und 'Treatment-

faktor'. Dabei werden additive und lineare Beziehungen formal angenommen.

MAIER & SELIGMAN (1976, S. 6) hingegen definieren die K-Gruppe als 'Kontrollgruppe', obwohl, verglichen mit der N-Gruppe, in der mathematischen Formalisierung sowohl für die K-Gruppe als auch für die UK-Gruppe ein 'Treatmentfaktor' hinzukommt.

Diese methodisch-statistische Indifferenz spiegelt sich auch in der Tatsache, daß es uns nicht gelang, in irgendeiner Veröffentlichung im Zusammenhang mit dem Namen SELIGMAN eine Begründung dafür zu entdecken, warum die N-Gruppe überhaupt verwendet wurde. Der erste Eindruck, das triadische Design sei simpel und elegant (SELIGMAN, 1975), bestätigt sich demnach nicht.

Es empfiehlt sich deshalb, nicht nur nach der Belastetheit der Theorie zu fragen, wenn wir im folgenden den empirischen Gehalt darstellen, sondern auch nach Hinweisen zu fahnden, die o. a. Indifferenz aufzuklären gestatten.

Das bislang Dargestellte läßt sich wie folgt zusammenfassen:

In der LH-Theorie wird angenommen, daß das Verhältnis von Reaktion und Ergebnis zu einer - zur objektiven Struktur des Reaktionskontingenzraums analogen - kognitiven Repräsentation führt. Folgt ein Ergebnis (Verstärker) mit gleicher Wahrscheinlichkeit auf eine Reaktion (R) wie auf ihre Unterdrückung ($\bar{R}$),

dann ist das Ergebnis unkontrollierbar und der Organismus hilflos, und diese Hilflosigkeit wird gelernt. Dieser aktive Lernprozeß ist beobachtbar am Eintreten eines LH-Effektes mit einem kognitiven, motivationalen und emotionalen Defizit.

Historisch gesehen wurde die LH-Theorie zunächst mit Tierexperimenten überprüft und erst später auf den Humanbereich ausgedehnt. Dabei zeigte sich eine Weiterentwicklung, die in den Tierexperimenten noch nicht gegeben ist. Wir haben Tier- und Humanexperimente deshalb getrennt dargestellt.

2.2.1.1.4.1 Darstellung eines Beispiels

MAIER & SELIGMAN (1976) beschreiben das LH-Phänomen in der experimentellen Situation wie folgt:

"Wenn man einen in bezug auf Experimente naiven Hund in eine Schüttelbox setzt, dann rast er bei Beginn des ersten elektrischen Schocks so lange wie ein Wahnsinniger umher, bis er zufällig über die Barriere springt und dem Schock entkommt. Beim nächsten Durchgang überquert der umherrasende Hund die Barriere schneller als beim vorhergehenden. Innerhalb weniger Versuchsdurchgänge lernt das Tier eine sehr effektive Fluchtreaktion, und es lernt bald, den Schock von vornherein zu vermeiden. Nach ungefähr fünfzig Durchgängen wird der Hund ruhig und gelassen vor der Barriere stehen. Bei Beginn des Ankündigungssignals für den Schock tut der Hund einen anmutigen Sprung und erhält nur noch selten Schocks. Ein ganz anderes Verhalten zeigen Hunde, die zuerst unvermeidbare Schocks in einer Hängematte erhalten hatten. Die ersten Reaktionen auf einen Schock sind wie die eines naiven Hundes. Er rast wie ein Wahnsinniger für ungefähr 30 Sekunden umher, dann bleibt er jedoch stehen, legt sich nieder und wimmert leise. Nach einer Minute endet der Schock automatisch. Der Hund springt nicht über die Barriere und entflieht

nicht dem Schock. Dies tut er auch nicht beim nächsten Durchgang. Zuerst strampelt er ein bißchen, scheint nach ein paar Sekunden aufzugeben und den Schock passiv hinzunehmen. Bei keinem der folgenden Durchgänge zeigt der Hund eine Fluchtreaktion." (MAIER & SELIGMAN, 1976, S. 4, übersetzt vom Verfasser)

SELIGMAN (1975) schreibt:

"Hier ist das typische Vorgehen, das wir zur Herstellung und zur Entdeckung der Gelernten Hilflosigkeit bei Hunden verwendeten. Am ersten Tag banden wir das Versuchstier in der Hängematte fest (the subject was strapped into the hammock) und gaben ihm 64 unvermeidbare elektrische Schocks, deren Zeitdauer fünf Sekunden und deren Intensität sechs Milliampere - mittelmäßig schmerzhaft - betrug. Den Schocks ging kein Signalreiz voraus, und sie wurden zeitlich zufällig gegeben. Vierundzwanzig Stunden später wurden mit dem Versuchstier zehn Versuchsdurchgänge eines Flucht-Vermeidungstrainings mit Ankündigungssignal in einer Zweiwegschüttelbox (two-way shuttle box) durchgeführt: Der Hund mußte über eine Barriere von einem Boxteil in den anderen springen, um dem Schock zu entfliehen oder ihn zu vermeiden. Die Schocks traten in beiden Boxteilen auf, so daß es keinen sicheren Ort gab, sondern allein die Reaktion des Springens immer zur Sicherheit führte. Mit dem Beginn eines Signals - Lichtflackern - begann jeder Versuchsdurchgang, und es wurde mit dem Ende des Versuchsdurchgangs beendet. Das Zeitintervall zwischen Signal und Schock betrug zehn Sekunden; wenn der Hund während dieses Intervalls die schulterhohe Barriere übersprang, endete das Signal, und der Schock fiel aus. Wenn der Hund während des Signal-Schock-Intervalls nicht sprang, erhielt er so lange einen 4,5-Milliampere-Schock, bis er die Barriere übersprang. Tat er dies nicht innerhalb von sechzig Sekunden nach Signalbeginn, endete der Versuchsdurchgang automatisch.

Zwischen 1965 und 1969 studierten wir das Verhalten von ungefähr 150 Hunden, die unvermeidbare Schocks erhalten hatten. Zwei Drittel davon - ungefähr 100 - waren hilflos. Die Tiere durchliefen die beeindruckende Aufgabensequenz, die ich beschrieben habe. Das

andere Drittel verhielt sich völlig normal; wie naive Hunde flohen sie und lernten bald, den Schock durch den rechtzeitigen Sprung über die Barriere zu vermeiden. Gelegentlich übersprangen hilflose Hunde zwischen den Versuchsdurchgängen die Barriere. Weiter: Wenn ein Hund in der linken Seite der Box gesessen und Schock nach Schock bekommen hatte, passierte es, daß der Hund - bei Öffnung der Tür auf der rechten Seite nach Beendigung des Versuchs - losschoß, um der Box als ganzer zu entfliehen. Da hilflose Hunde physisch in der Lage waren, die Barriere zu überspringen, muß ihr Problem psychologischer Art gewesen sein." (SELIGMAN, 1975, S. 23f., Übersetzt vom Verfasser)

In beiden Zitaten wird das Verhalten der Tiere in der UK-Gruppe ausführlich dargestellt. Wir erfahren nichts über Unterschiede zwischen K- und N-Gruppe.

Immerhin kann man obige Ergebnisse nicht als Artefakt abtun. Sie wurden an

- Hunden (OVERMIER, 1968; OVERMIER & SELIGMAN, 1967; SELIGMAN & GROVES, 1970; SELIGMAN & MAIER, 1967; SELIGMAN, MAIER & GEER, 1968),
- Katzen (THOMAS & BALTER, zit. n. MAIER & SELIGMAN, 1976; vgl. auch: MASSERMAN, 1943; SEWARD & HUMPHREY, 1967; ZIELINSKI & SOLTSYK, 1964),
- Fischen, insbesondere bei Goldfischen (PADILLA et al., 1970; vgl. PADILLA, 1973; BEHREND & BITTERMAN, 1963; BINTZ, 1971; FRUMKIN & BROOKSHIRE, 1969; PICKNEY, 1967),
- Ratten (MAIER, ALBIN & TESTA, 1973; MAIER & TESTA, 1975; SELIGMAN & BEAGLEY, 1975; SELIGMAN ROSSELINI & KOZAK, 1975; WEISS et al., 1968;

vgl. die Übersichten bei SELIGMAN et al., 1971; MAIER, SELIGMAN & SOLOMON, 1969; Analogien bei McCULLOCH & BRUNER, 1939; MOWRER & VIEK, 1948),

- Mäusen (BRAUD, WEPMAN & RUSSO, 1969),
- Schaben (vgl. HORRIDGE, 1962) und
- Tauben (vgl. ENGBERT, HANSEN, WELKER & THOMAS, 1973)

repliziert bzw. es wurden entsprechende Defizite berichtet. Auf die Ergebnisse von Humanexperimenten werden wir weiter unten eingehen.

Es zeigt sich, daß die motivationalen, kognitiven und emotionalen Aspekte des LH-Effektes für viele Species replizierbar sind.

Die Frage stellt sich, inwieweit der im triadischen Design erzielte LH-Effekt ein Laborartefakt darstellt und inwieweit der LH-Effekt sich auch außerhalb der experimentellen Situation zeigt.

2.2.1.1.4.2 Transfer der Gelernten Hilflosigkeit

Auch hier wollen wir uns zunächst auf die Darstellung der Ergebnisse aus Tierexperimenten beschränken.

OVERMIER & SELIGMAN (1967) und OVERMIER (1968) sowie SELIGMAN & GROVES (1970) variierten die Zeitdauer der Unterbrechung im triadischen Design. Danach reagieren Hunde, die in einer Sitzung mit unkontrollierbaren Schocks behandelt wurden, nur dann hilflos,

wenn die Testphase binnen vierundzwanzig Stunden folgt. Nach achtundvierzig Stunden ist der LH-Effekt 'verschwunden', d. h. die Tiere reagieren, als hätten sie diese Vorerfahrung nie gemacht und übertragen den LH-Effekt nicht auf neue Problemsituationen.

Man kann deshalb annehmen, daß bei Hunden ein schwacher LH-Effekt einem zeitlichen Verlauf (time-course) unterliegt, sich die 'unkontrollierbare' Gedächtnisspur nach einiger Zeit auflöst.

MAIER & SELIGMAN (1976, S. 14f.) betonen, daß ein zeitlicher Verlauf nur einmal beobachtet werden konnte: bei Hunden unbekannter Vorgeschichte, die eine Schocksitzung erhielten. SELIGMAN & GROVES (1970) hingegen fanden bei (isoliert aufgezogenen) Hunden einen LH-Effekt noch nach einer Woche.

Bei Ratten scheint das Entstehen des LH-Effektes von der Schwierigkeit der Kontrollreaktion abzuhängen: Bei einfachen Reaktionen ergibt sich überhaupt kein LH-Effekt, bei schwierigen Reaktionen ein chronischer (vgl. MAIER & SELIGMAN, 1976, S. 8f.; SELIGMAN, MAIER & SOLOMON, 1971).

Abgesehen von der Diskussion des Zeitverlaufs ('time course') finden wir in der Literatur nur sporadisch und anekdotisch Hinweise auf einen Transfer des LH-Trainings.

SELIGMAN (1975, S. 24f.) beschreibt in einem kurzen Abschnitt das Verhalten hilfloser Hunde. Diese

scheinen sich demnach nur so lange passiv zu verhalten, wie eine Flucht aus dem Versuchsaufbau unmöglich ist (vgl. auch SELIGMAN, S. 119). Andererseits heißt es:

"Außerhalb der Schüttelbox unterscheiden sich hilflose Hunde in ihrem Verhalten ebenfalls von nicht-hilflosen. Wenn ein Versuchsleiter zum Heimatkäfig (home-cage) geht, um einen nicht-hilflosen Hund zu holen, dann fügt sich dieser nicht willffähig: er bellt, rennt zur Rückwand und sträubt sich. Im Gegensatz dazu scheinen hilflose Hunde willig zu sein (seem to wilt); sie sinken passiv auf den Boden, rollen sich gelegentlich sogar auf den Rücken und nehmen eine Demutshaltung ein; sie zeigen keinen Widerstand." (SELIGMAN, 1975, S. 25, übersetzt vom Verfasser)

SELIGMAN (1974, S. 117) erwähnt beiläufig, daß Ratten nach unkontrollierbaren Schocks weniger Schmerz-aggression ('shock elicited aggression') zeigen. Ratten, die mit kontrollierbaren Schocks behandelt wurden, verhalten sich dagegen normal: sie attackieren eine andere Ratte des gleichen Geschlechts, wenn ihnen in deren Anwesenheit Schmerzen zugefügt werden.

Ferner soll sich ein LH-Experiment auf die soziale Hierarchie auswirken: Ratten aus der UK-Gruppe sinken in der Hierarchie, Ratten aus der K-Gruppe steigen auf.

SELIGMAN schließt daraus, daß Dominanz und Aggression der Kontrollierbarkeit von Schocks zugeordnet werden können.

Ein weiteres anekdotisches Ergebnis betrifft das

Verhalten junger Hunde (SELIGMAN, 1974, S. 124). Zwölf Würfe junger Welpen wurden unmittelbar nach der Entwöhnung in drei Gruppen geteilt und unkontrollierbaren, kontrollierbaren oder keinen Schocks ausgesetzt. Danach wartete der Forscher zwölf Monate, um die Hunde dann verschiedenen Versuchen zu unterziehen. In Lernexperimenten zeigten sich keine Gruppenunterschiede, wohl aber in der sozialen Hierarchie.

Werden Ratten nach der Entwöhnung der Pretestphase des triadischen Designs unterworfen, so zeigt die UK-Gruppe noch im Erwachsenenalter den LH-Effekt (SELIGMAN, 1975, S. 148).

Dies läßt sich nicht nur für unkontrollierbare Schocks nachweisen, sondern auch für unlösbare (für die K-Gruppe lösbare) Diskriminanzaufgaben (vgl. BAINBRIDGE, 1973).

Vom Non-plus-ultra eines chronischen LH-Effektes berichtet SELIGMAN (1975, S. 149) in aller Kürze: Werden weibliche Ratten vor ihrer Befruchtung mit kontrollierbaren, unkontrollierbaren oder keinen Schocks behandelt, dann kann noch bei den erwachsenen Nachkommen der UK-Gruppe ein LH-Effekt beobachtet werden!

Insgesamt gesehen sind die Ergebnisse der tierexperimentellen Forschung zum Transfer des LH-Effekts teilweise inkonsistent.

Bei Hunden konnte ein zeitlicher Verlauf des LH-

Effektes beobachtet werden, ebenso wie ein chronischer LH-Effekt erzeugt werden konnte. Bei einer Trainingssitzung 'verschwindet' der Effekt nach einiger Zeit, bei mehreren Sitzungen wird er chronisch; nach einem Jahr zeigt sich dann kein motivationales, kognitives oder emotionales Defizit mehr, wohl aber eine Veränderung in der sozialen Hierarchie.

Bei Ratten tritt ein LH-Effekt entweder erst gar nicht auf oder er zeigt sich sofort als chronisch-resistent. Möglicherweise wird der Effekt dann über mehrere Generationen hinweg weitergegeben. Für letzteres finden wir als Beleg aber nur SELIGMANS anekdotische Beschreibung. Wir können nur spekulieren, wie ein solcher Generationseffekt zustande kommen kann; wir können spekulieren, daß zur Entstehung eines LH-Effektes eine direkte individuelle Erfahrung mit unkontrollierten Ergebnissen gar nicht notwendig ist, sondern daß ein hilfloses Modell ausreicht.

2.2.1.2 Kritische Würdigung der Phase I

Auf S. 73 haben wir zitierend beschrieben, daß Hunde u. U. nur so lange ein Flucht- und Vermeidungsdefizit (ein motivationales Defizit) zeigen, wie ihnen die Flucht aus der gesamten Experimentierbox verwehrt wird. Zwar sollen sie sich auch außerhalb der experimentellen Situation anders verhalten, dies wird jedoch für folgenden Fall beschrieben: "When an experimenter goes to the home-cage..."

(SELIGMAN, 1975, S. 25). Ein 'experimenter' - ein Versuchsleiter also - ist aus der Sicht eines Hundes jedoch ein Stimulus, der zum experimentellen Setting zählt.

Wir stellen uns daher die Frage, ob nicht die gesamte experimentelle Situation zum diskriminativen Reiz oder zum bedingten Reiz für hilfloses Verhalten wurde.

SELIGMAN (1975, S. 60) selbst diskutiert im Zusammenhang mit dem Immunisierungseffekt (der Umkehrung des LH-Effektes) das Problem der diskriminativen Kontrolle: Eine Person kann lernen, in der Situation A keine Kontrolle zu haben, wohl aber in der Situation B. Beispiele hierfür bieten bislang nur Humanexperimente (DWECK & REPUCCI, 1973; ROTH & BOOTZIN, 1974). Im Tierexperiment ist es bislang noch nicht gelungen, hilfloses Verhalten unter diskriminative Kontrolle zu bringen.

Erst dann wäre evident, daß der LH-Effekt ausschließlich und notwendig eine Folge der Konfrontation mit unkontrollierbaren Verstärkern ist. Bislang ist offen, ob nicht gleichzeitig eine emotionale Reaktion auf einen zufällig auftretenden oder gar unvermeidbaren Stimulus (z. B. Versuchslabor, Versuchsleiter etc.) konditioniert wurde, ob nicht eine Reiz-Reaktions-Kontiguität verhaltensrelevanter ist als eine Reaktions-Verstärker-Kontingenz (vgl. LEVIS, 1976; BLACK, 1977). Die LH-Forscher diskutieren die konkurrierenden Erklärungsansätze ausführlich und weisen jeden Alternativansatz zurück (SELIGMAN, 1975; MAIER & SELIGMAN, 1976; MAIER, 1977).

Der theoretische Streit ist noch nicht abgeschlossen, wenngleich wir den Eindruck haben, daß die LH-Forscher das Modell mit dem größten Integrationswert vorlegen.

Ein Blick auf die empirischen Daten (vgl. S. 72) zeigt, daß es sich beim LH-Effekt mit Sicherheit um kein Zufallsphänomen handelt. Wie sich bei der Diskussion der Humanexperimente auch noch zeigen wird, tritt er - bei Beachtung bestimmter experimenteller Durchführungsrichtlinien - reliabel auf.

Die LH-Theorie hat jedoch einige unübersehbare Lücken:

- Die Theorie der LH macht keine Aussage über das Verhältnis von K- und N-Gruppe. Wir sind uns nicht sicher, ob aus der Bemerkung bei MAIER & SELIGMAN (1976, S. 18), die Reaktions-Ergebnis-Unabhängigkeit werde aktiv gelernt und sie hemme dann proaktiv durch assoziative Interferenz die Ausbildung entgegengesetzter Formen des Lernens, und aus der wenige Zeilen später zu lesenden Argumentation, ein naiver Hund unterscheide sich von einem 'hilflosen' dadurch, daß er keine interferierende Erwartung habe, ob aus diesen beiden Komponenten abgeleitet werden darf, K- und N-Gruppe würden sich, weil sie beide keine interferierenden Erwartungen hätten, nicht unterscheiden. Ein Indiz dafür liefern HIROTO & SELIGMAN (1975), wenn sie beim Vergleich der UK-Gruppe mit den beiden anderen Gruppen einen einseitigen statistischen Test, beim Vergleich von K- und N-Gruppe einen zweiseitigen verwenden.
- Bei Untersuchungen mit dem triadischen Design blieb bislang offen, ob - insbesondere bei der UK-Gruppe - sich konditionierte emotionale Reaktionen herausbilden, ob diese respondenten Reaktionen mit anderen Reaktionen interferieren, ob sie unter operante Kontrolle (z. B. durch den Versuchsleiter) kommen, ob die Reaktionen unter dis-

kriminative Kontrolle gebracht werden können. Dies stört um so mehr, als SELIGMAN (1975, S. 12) die Situation des Versuchstieres beim Klassischen Konditionieren mit Hilflosigkeit gleichsetzt. Wir fanden kein Experiment, in dem das Treatment der UK-Gruppe - also potentiell Klassisches Konditionieren - durch ein explizites Klassisches Konditionieren ersetzt wurde, wohl aber zwei Bemerkungen (MAIER & SELIGMAN, 1976, S. 7; SELIGMAN, 1975, S. 24), daß dies den LH-Effekt nicht verändert.

- Sowohl das Problem der diskriminativen Kontrolle wie das Problem der Hilflosigkeit beim Klassischen Konditionieren werden berührt, wenn man die LH-Theorie in Beziehung zur Zwei-Stufen-Theorie der Entstehung der Phobie nach MOWRER (1939) setzt. Letztere Theorie sieht vor, daß die Genese einer phobischen Reaktion in zwei Stufen erfolgt: In der ersten Stufe wird durch Stimulussubstitution Angst an einen vorher neutralen Reiz gekoppelt, in der zweiten Stufe wird jede Vermeidungsreaktion negativ verstärkt, die die Angst reduziert. LH-theoretisch lernt das Tier in der ersten Stufe Hilflosigkeit, was dann dazu führen kann, daß wegen des kognitiven Defizits in der zweiten Stufe keine operante Reaktion - also auch keine Vermeidungsreaktion - mehr gelernt wird.
- Die Durchführung der Experimente bleibt in manchen Punkten unklar. Es heißt lediglich, daß während der Pretreatmentphase K-Gruppe und UK-Gruppe gekettet ('yoked') sind, wobei der UK-Gruppe Kontrolle verwehrt wird. Man muß annehmen, daß jeweils Paare von Versuchstieren in je einen Versuchskäfig gebracht werden und während des Experiments dann beide in Abhängigkeit von den Reaktionen des einen geschockt werden oder daß der Schock automatisch gesteuert wird. In den von uns zitierten Beispielen trifft letzteres zu; jedoch gibt es eine Reihe von Studien, in denen der Schock handgesteuert gegeben worden sein muß. Es bleibt unklar, ob die Tiere einander sehen konnten oder ob der Versuchleiter die Tiere der K-Gruppe intensiv beobachtete und die Tiere der UK-Gruppe nicht beachtete. Es bleibt offen, ob außer dem expliziten Treatment noch andere Variablen (z. B. ein HAWTHORNE-Effekt) wirkten oder ob sich die Experimentatoren dagegen absicherten.

Zusammenfassend läßt sich zu den Ergebnissen der Tierversuche im triadischen Design sagen, daß der LH-Effekt mehrfach überzeugend repliziert wurde und kein Zufallsphänomen darstellt. Nicht ganz überzeugend sind die Ergebnisse zum zeitlichen Verlauf und zum Transfer auf Problemsituationen außerhalb des Labors, wenngleich sie als Hinweis für die Annahme ausreichen, daß hilflose Tiere in der Schüttelbox mehr lernen als die Kontingenzen des Reaktionskontingenzraumes.

So überzeugend die empirischen Ergebnisse auch sein mögen, so benötigen wir für die Interpretation der Ergebnisse, die mit dem triadischen Design gewonnen werden, nicht notwendig das theoretische Element der kognitiven Repräsentation, sondern wir können auf eine mehrfach replizierte objektive Wenn-Dann-Reaktion verweisen: wenn objektive Unkontrollierbarkeit vorliegt, dann zeigt sich in einer zeitlich späteren Phase des Experiments ein LH-Effekt. Erst das Immunisierungssparadigma und das Konzept der Vorerfahrung verweisen zwingend auf kognitive Zwischenprozesse.

An späterer Stelle werden wir das LH-Experiment aus der Sicht von Versuchspersonen ausführlich reinterpretieren. Im Tierexperiment ist es so, daß sich das Versuchstier zu Beginn des Experiments in einer Hängematte mit eingeschränkten Verhaltensmöglichkeiten befindet und einen Schock bekommt (Ausgangssituation). Dieser ist so intensiv, daß die Beendigung des Schocks vom Tier angestrebt wird (Zielzustand). Aufgabe des Tieres ist es nun, Verhaltensweisen zu

finden, die Ausgangssituation und Zielzustand verknüpfen. Bei der UK-Gruppe wird dies vom Versuchsleiter systematisch verhindert.

In einem zweiten Teil befinden sich die Tiere z. B. in einer Schüttelbox mit ebenfalls eingeschränkten Verhaltensmöglichkeiten und erhalten wiederum einen Schock, der so intensiv ist, daß seine Beendigung angestrebt wird. Aufgabe des Tieres ist es wiederum, Operationen zu suchen und zu realisieren, die Ausgangssituation und Zielzustand verknüpfen. Dabei zeigt sich jedoch, daß - bei objektiv gleichem Zielzustand und objektiv gleicher Ausgangssituation - sich die UK-Gruppe beim Problemlösen anders verhält als K- und N-Gruppe: die kognitive Repräsentation der 'unkontrollierbaren' Bedingung beeinflusst offenbar den Suchprozeß.

Die Tiere der UK-Gruppe sind unfähig, Ausgangssituation und Zielzustand mit Operationen zu verknüpfen. Dies beschreiben die Hilfflosigkeitsforscher mit dem Terminus 'kognitives Defizit'.

Die Beeinflussung eines Suchprozesses beim Problemlösen hatten wir in einem früheren Kapitel als Strategie definiert. Demnach haben die 'hilfflosen' Tiere eine Strategie gelernt.

Wir wollen diesen Punkt am Beispiel der Humanexperimente näher diskutieren und gehen deshalb zur Darstellung der Phase II der Entwicklung der LH-Theorie über, die mit dem Beginn der Humanexperimente zusammenfällt.

2.2.2 Phase II: Humanversuche -

Betonung der Kognitiven Komponente

Bis hierhin haben wir die Manipulation der Vorerfahrungen im triadischen Design des Tierexperiments herausgearbeitet. Die Diskussion der Humanversuche wollen wir wieder mit einem Beispiel einleiten.

2.2.2.1 Darstellung eines Beispiels

HIROTO (1974) wählte 96 Versuchspersonen nach den Ergebnissen eines Fragebogens zum Locus-of-Control aus.

Das Konzept des Locus-of-Control kann man frei mit 'Kontrollorientierung' übersetzen. Es stammt aus der von ROTTER (1954) vorgelegten Social-Learning-Theorie. Übersichtsreferate wurden von LEFCOURT (1966), ROTTER (1966) und JOE (1971) vorgelegt.

"In Rotter's social learning theory ... the potential for any behavior to occur in a given situation is a function of the person's expectancy that the given behavior will secure the available reinforcement, and the value of the available reinforcement for that person ... the control construct is considered a generalized expectancy, operating across a large number of situations, which relates to whether or not the individual possesses or lacks power over what happens to him... internal control refers to the perception of positive and/or negative events as being a consequence of one's own actions and thereby under personal control; external control refers to the perception of positive and/or negative events as being unrelated to one's own behaviors in certain situations and therefore beyond personal control." (LEFCOURT, 1966, S. 207)

Wir sprechen demnach von internaler Kontrolle, wenn Ereignisse in Abhängigkeit von den eigenen Aktionen gesehen werden, und von externaler Kontrolle, wenn Ereignisse und Aktionen subjektiv unabhängig sind.

HIFOTO wählte 96 Versuchspersonen nach ihrer Kontrollorientierung aus und verteilte sie auf die drei Treatmentbedingungen des triadischen Designs. Innerhalb jeder Treatmentbedingung teilte er die Versuchspersonen in je zwei Gruppen mit unterschiedlicher Instruktion.

Der Versuch begann für die UK- und die K-Gruppen mit der Pretestphase, für die N-Gruppen mit der Testphase.

In der Pretestphase hatten alle teilnehmenden Versuchspersonen ein Brett mit einem roten Knopf vor sich und erfuhren über ein Tonband, daß ein lauter, aversiver Ton auftreten würde und daß sie eine Möglichkeit hätten, diesen zu beenden. Sie erhielten keine weiteren Informationen. Danach erhielten alle Versuchspersonen über einen Kopfhörer eine Serie von Tönen (3000 Hz, 5 Sekunden, 110 dB). Die einzige Bedingungsvariation in dieser Phase war, daß die Versuchspersonen der K-Gruppen die Möglichkeit hatten, durch einen Druck auf den roten Knopf den Ton vorzeitig zu beenden (Fluchtreaktion), während für die Versuchspersonen der UK-Gruppe der rote Knopf keine Schaltfunktion hatte.

In der Testphase, die im selben Raum vom selben Versuchsleiter durchgeführt wurde, hatten die Versuchspersonen dann ein Brett mit einem Hebel vor sich. Sie erfuhren über ein Tonband, daß eine weitere Serie von aversiven Tönen gegeben würde und daß es wieder eine Möglichkeit gebe, den Ton zu beenden.

Im Anschluß an diesen gemeinsamen Teil wurde der Hälfte der Versuchspersonen eine 'skill'-Instruktion überspielt, die wir übersetzt haben:

"Was Sie tun können, müssen ganz allein Sie herausfinden. Es gibt eine Lösung des Problems. Wenn Sie sie finden, hört der Ton auf. Das Ausmaß des Unangenehmen, das Ihnen begegnet, ist deshalb abhängig von Ihren Fähigkeiten und Fertigkeiten, das Problem zu lösen. Sie haben die Chance, die Situation zu kontrollieren." (HIROTO, 1974, S. 189, übersetzt vom Verfasser)

Der anderen Hälfte der Versuchspersonen wurde folgende Instruktion vorgespielt ('chance'-Instruktion):

"Ich (HIROTO) für meinen Teil werde die Lösung des Problems kontrollieren. Mit anderen Worten; die technische Möglichkeit, den Ton zu stoppen, liegt ganz allein bei mir. So weit, wie Sie betroffen sind, handelt es sich um ein Ratespiel. Wenn Sie richtig raten, hört der Ton automatisch auf. Wenn Sie aber falsch raten, ertönt er weiter." (HIROTO, 1974, S. 189, übersetzt vom Verfasser)

Der Instruktion folgte eine Serie von Licht-Ton-Kombinationen: Fünf Sekunden vor dem aversiven Ton leuchtete eine rote Lampe auf und erlosch zusammen mit diesem. In dieser Phase hatten alle Versuchspersonen die Möglichkeit, durch Bewegen eines Hebels eine Stoppschaltung zu steuern, also eine Vermeidungsreaktion (Stopp vor Tonbeginn) oder eine Fluchtreaktion (Stopp nach Tonbeginn) auszuführen.

Als abhängige Variablen wurden in der Testphase u. a. die Anzahl der Vermeidungsreaktionen, die Anzahl der Durchgänge ohne Fluchtreaktion und die gemittelte Reaktionslatenz für alle Durchgänge gemessen. HIROTO

gibt an, dies seien ähnliche Indikatoren, wie sie in den Tierexperimenten verwendet wurden.

Im folgenden nennen wir die wichtigsten Ereignisse:

Das Rating der Aversivität des Ton ergab keine Gruppenunterschiede.

Die Versuchspersonen der UK-Gruppen realisierten in über 50 % der Durchgänge keine Fluchtreaktion, während die Versuchspersonen der K-Gruppen nur in 13 %, die der N-Gruppen nur in 11 % der Durchgänge nicht reagierten. Die statistische Analyse ergab, daß sich die UK-Gruppen von den K- und den N-Gruppen signifikant unterscheiden, letztere unterschieden sich nicht. HIROTO interpretiert dies als motivationales Defizit.

Versuchspersonen mit externaler Kontrollorientierung zeigten - unabhängig vom Pretreatment - längere Reaktionslatenzen als die Versuchspersonen mit internaler Kontrollorientierung und erreichen eine geringere Zahl von Vermeidungsreaktionen. HIROTO meint dazu, daß der Glaube an die Unabhängigkeit von Reaktion und Ergebnis (externale Kontrollorientierung) die gleiche Wirkung hat wie objektive Unkontrollierbarkeit.

Die 'chance'-instruierten Gruppen zeigten - unabhängig von Pretreatment - längere Reaktionslatenzen und erreichten eine geringere Zahl von Vermeidungsreaktionen als die 'skill'-instruierten. HIROTO interpretiert dies als gleichsinnige Wirkung von Erfahrung (der objektiven Unkontrollierbarkeit) und ('chance'-) Instruktion.

Der Autor fand keine Interaktion zwischen den Treatmentbedingungen und/oder den Instruktionen und/oder der Kontrollorientierung. Er berichtet nichts über emotionale Folgen des Experiments. Die durchweg fehlenden Unterschiede zwischen N- und K-Gruppen greift er nicht auf.

Dafür finden wir bei HIROTO einen Ansatz, der aus den Tierexperimenten her unbekannt ist. Im Locus-of-Control-Konzept ist eine Versuchsperson mit externa-

ler Kontrollorientierung jemand, der positive und/oder negative Ergebnisse unabhängig von seinem Verhalten und daher nicht unter seiner Kontrolle sieht; ein Mensch mit internaler Kontrollorientierung sieht die Ergebnisse vorwiegend in Abhängigkeit von seinen Reaktionen und daher unter seiner Kontrolle (vgl. ROTTER, 1966; LEFCOURT, 1966; JOE, 1971).

Innerhalb der Locus-of-Control-Forschung können wir zwei methodische Ansätze unterscheiden. Zum einen wird die Kontrollorientierung mit diversen Fragebögen erfaßt und mit anderen Fragebogendaten korreliert; die Mehrheit der Veröffentlichungen besteht aus Korrelationsstudien, eine starke Minderheit verwendet die Kontrollorientierung aber auch als organismische Variable in faktoriellen Designs. Zum anderen wird die Kontrollorientierung durch 'skill'- und 'chance'-Instruktionen manipuliert und deren Auswirkungen auf abhängige Variablen als Treatmentfaktor im faktoriellen Design erfaßt.

HIROTO (1974) verwendet beide methodischen Ansätze. Er kommt zu der Schlußfolgerung, daß dem LH-Treatment im triadischen Design, der externalen Kontrollorientierung und der 'chance'-Instruktion etwas Gemeinsames zukommt: der Erwartung der Unabhängigkeit von Reaktion und Verstärkung. Die zwar nicht identischen, aber gleichsinnigen Ergebnisse seines Experiments lassen diese Folgerung zu.

Nicht unbedingt die Messung des Locus-of-Control, wohl aber eine direkte Messung der Erwartung hatten wir bereits als Möglichkeit angedeutet (vgl. S. 65)

HIROTO verweist auf eine weitere Möglichkeit, die Kontingenz-Erwartung zu manipulieren: das Medium der Instruktion, einer gezielten Kommunikation.

HIROTOS Ergebnisse haben erhebliche Konsequenzen für die LH-Theorie bzw. für ihre Darstellung. Weiter oben haben wir geschrieben, daß die in den Überblicksarbeiten dargestellte enge Koppelung zwischen Reaktionskontingenzzraum und LH-Effekt möglicherweise irreführend ist. HIROTO nun weist nach, daß geeignete Instruktionen und eine allgemeine Kontrollorientierung zu sehr ähnlichen Ergebnissen führen wie objektiv unkontrollierbare Bedingungen. Dies bestätigt unsere Schlußfolgerung, daß die zentrale Komponente der LH-Theorie die kognitive Repräsentation ist, gleichgültig wie sie entstanden sein mag. Es kommt demnach lediglich darauf an, daß die kognitive Repräsentation - SELIGMAN spricht von 'Erwartung' - der Unabhängigkeit von Reaktionen und Verstärkern wirksam ist. Dies ist eine Weiterentwicklung der Theorie, die sich in Phase II findet.

2.2.2.2 Transfer der Gelernten Hilflosigkeit

HIROTO & SELIGMAN (1975) stellten sich die Frage, ob der LH-Effekt in der Testphase nur bei ähnlichen Aufgaben oder auch bei sehr unähnlichen auftritt.

Sie trainierten in vier unabhängigen Experimenten ihre Versuchspersonen in der Pretestphase entweder

- mit dem o. a. HIROTO-Procedere (instrumentelle Bedingung) oder

- führten ein Experiment zur Konzeptbildung durch (kognitive Bedingung).

In der Testphase wurden je zur Hälfte entweder

- der HIROTO-Testphasenaufbau verwendet (instrumentelle Bedingung) oder
- sie ließen die Versuchspersonen Anagramme lösen (kognitive Bedingung).

Es entstanden so die Treatmentkombinationen der vier Experimente:

1. instrumenteller Pretest - instrumenteller Test,
2. instrumenteller Pretest - kognitiver Test,
3. kognitiver Pretest - instrumenteller Test,
4. kognitiver Pretest - kognitiver Test!

Zur Konzeptbildung in der Pretestphase der Experimente drei und vier wurden den Versuchspersonen mehrmals je zwei vierdimensionale Stimuli gezeigt. Jede der Dimensionen wurde durch zwei Ausprägungen repräsentiert:

1. Buchstabe (A oder T),
2. Größe des Buchstabens (groß oder klein),
3. Farbe der Buchstaben (vollschwarz oder weiß mit Umrandung),
4. Umrandung der Buchstaben (Kreis oder Quadrat).

Die beiden simultan dargebotenen Stimuli wurden in allen Dimensionen mit dem jeweils anderen Wert gegeben: ein großes, schwarzes A im Kreis wird zusammen mit einem kleinen, weißen T im Quadrat gezeigt. Bei HIROTO & SELIGMAN (1975) wurde jeweils einer der 4 x 2-Werte (für die K-Gruppen) als richtig definiert, z. B. quadratische Umrandung, den Versuchspersonen aber nicht mitgeteilt. Die Versuchspersonen mußten im Verlauf der Darbietung mehrerer Stimulipaare herausfinden, welcher Wert 'richtig' ist. Dabei zeigten sie auf einen der Stimuli, und der Versuchsleiter teilte ihnen das Ergebnis mit.

Durch Exklusion falscher Werte finden sie das richtige Konzept.

Bei der UK-Gruppe wurde ein vorher festgelegtes Antwortschema gegeben; am Ende des Versuchs wurde ein dann geratener Wert immer mit der Antwort 'falsch' beantwortet.

Den N-Gruppen wurden die Stimuli ohne Erläuterung gezeigt.

HIROTO & SELIGMAN betonen, daß unlösbare kognitive Aufgaben eine formale Analogie für Unkontrollierbarkeit darstellen, "... since in both the probability of reinforcement (correct or incorrect, or shock or no shock) is independent of responding." (HIROTO & SELIGMAN, 1975, S. 313)

Die instrumentelle Pretestphase war die gleiche wie bei HIROTO (1974), nur daß bei dem jetzigen Versuch der N-Gruppen die Sequenz aversiver Töne, wie sie auch die UK-Gruppen erhielten, ohne nähere Erläuterung überspielt wurde. Die Instruktion beinhaltete den Hinweis, daß das Lauschen auf Töne Teil des Experiments sei. Die N-Gruppen wurden angewiesen, passiv zu bleiben und sich die Töne anzuhören.

Die Versuchspersonen der instrumentellen Treatmentgruppen wurden instruiert, den richtigen Weg, 'den Lärm zu stoppen', sei nicht, den Kopfhörer abzunehmen oder den Versuchsaufbau zu demontieren.

Die Autoren erwähnen in einem Satz, daß allen Versuchspersonen nach ein paar Probetönen die Gelegenheit gegeben wurde, den Versuch abzubrechen. Es wird nicht berichtet, daß eine Versuchsperson dies tat.

Die instrumentelle Testphase war die gleiche wie bei HIROTO (1974).

Die kognitive Testphase beinhaltete das Lösen von Anagrammen.

Anagramme sind Buchstabenfolgen, die entstehen, wenn die Buchstaben eines geläufigen Wortes in eine andere Reihenfolge gebracht werden. HIROTO & SELIGMAN bildeten alle Anagramme aus fünfziffrigen Wörtern und alle nach dem gleichen Schema: aus der Reihenfolge 1-2-3-4-5 wurde 3-4-2-5-1, sie konstruierten z. B. aus 'CLERK' 'ERLKC', aus 'FAULT' wurde 'ULATF' und aus 'HABIT' entstand 'BIATH'. Die Versuchspersonen wurden in der Instruktion darauf hingewiesen, daß es ein Muster oder ein Prinzip für die Lösung aller Anagramme gibt; ihnen wurde nicht mitgeteilt, was damit gemeint war.

Die Autoren wiesen den LH-Effekt in drei der vier Experimente auf Anhieb nach. Den LH-Effekt für das Experiment mit kognitivem Pretest - kognitivem Test konnten sie erst nachweisen, als sie in einem fünften Experiment die Zahl der Durchgänge im kognitiven Pretest von drei auf vier erhöhten.

In dem dargestellten Experiment tauchen erstmals einige Ergebnisse der Pretestphase auf. Danach war es für die Pretestphase ohne Einfluß, welche Testphase folgte (da die entsprechende Instruktion erst in der Unterbrechung gegeben wurde, wäre dies auch verwunderlich). Ferner berichten die Autoren, daß die Versuchspersonen der UK-Gruppen verglichen mit den K-Gruppen ihr Abschneiden stärker sowohl auf ihre Unfähigkeit als auch auf die Schwierigkeit der Aufgaben attribuieren.

HIROTO & SELIGMAN verrechnen als abhängige Variablen drei (nicht unabhängige) Kriterien: Kriterium 1 ist die mittlere Anzahl der Versuche bis zur Akquisition des Anagrammlösungsprinzips, Kriterium 2 ist die mittlere Anzahl ungelöster Anagramme, und Kriterium 3 ist die mittlere Reaktionslatenz. Wie aus Tabelle 1 sofort zu entnehmen ist, liefern die drei Kriteriumsvariablen nicht in jedem Fall identische Prüfstatistiken.

Aus den von den Autoren veröffentlichten Graphiken ist zu entnehmen, daß die K-Gruppen in allen Fällen kleinere Mittelwerte erreichen als die N-Gruppen und diese wiederum einen kleineren als die UK-Gruppen. Beim Vergleich der beiden Extremgruppen fallen - wie man in Tabelle 1 sieht - auch nur zwei von zwölf (einseitigen) t-tests nicht signifikant aus. Der (zweiseitige) Vergleich der K-Gruppen und der N-Gruppen fällt - obwohl der Mittelwert der K-Gruppen in allen zwölf Graphiken niedriger ist - stets nicht-signifikant aus. Beim (einseitigen) Vergleich von UK-Gruppen und N-Gruppen scheint es vom verrechneten Kriterium abzuhängen, ob wir die Nullhypothese annehmen oder zurückweisen müssen.

Die Analyse der Begleitfragebögen ergab, daß sich die Gruppen nicht hinsichtlich der Anstrengung ('amount of trying') unterschieden. Die UK-Gruppen fühlten sich signifikant 'frustrierter', insbesondere die UK-Gruppe mit kognitiver Pretestphase.

Tabelle 1: Zusammenfassung der Ergebnisse der Testphase von HIROTO & SELIGMAN (1975); berechnet wurden t-tests zwischen den Gruppen mit kontrollierbaren ('K') und unkontrollierbaren ('UK') Bedingungen und der Kontrollgruppe ('N'); weitere Erläuterungen im Text!

	K vs. UK			K vs. N			UK vs. N			Bemerkung
	Krit.1	Krit.2	Krit.3	Krit.1	Krit.2	Krit.3	Krit.1	Krit.2	Krit.3	
I str. vs. Instr.	p<.01 xx	p<.01 xx	p<.01 x	p>.05	p>.05	p>.05	p<.05 x	p>.05	p>.05	Replikation durch HIROTO(1974)
Instr. vs. Cogn.	p>.05	p<.02 x	p<.02 x	p>.05	p>.05	p>.05	p>.05	p<.05 x	p<.05 x	Replikation durch GATCHEL & PROCTOR (1975)
Cogn. vs. Instr.	p<.01 xx	p<.01 xx	p<.01 xx	p>.05	p>.05	p>.05	p<.05 x	p<.02 x	p>.05	
Cogn. vs. Cogn. Exp.	p<.02 x	p<.02 x	p>.05	p>.05	p>.05	p>.05	p<.05 x	p<.05 x	p>.05	diverse Replikationen

Die Autoren interpretieren ihre Ergebnisse als Nachweis für 'cross-modal helplessness' und verweisen darauf, daß obige Ergebnisse nicht mit der Annahme konkurrierender, interferierender Reaktionen oder eines behindernden organismischen Zustandes wie Ermüdung erklärt werden können. Der LH-Effekt trat unabhängig vom Operationsmodus (Knopfdrücken bzw. Konzeptbildern) auf.

"Our present findings provide the coup de grace against any peripheralist interpretation. No competing motor response could generate cognitive interference from instrumental inescapability or instrumental interference from cognitive insolubility." (HIROTO & SELIGMAN, 1975, S. 326)

Auch Versuche, ihre Ergebnisse mit Frustration, unterschiedlicher Aversivität des Tons und daraus ableitbaren response-cost-Kalkulationen oder mit Demand-Characteristics der Versuchssituation (vgl. ORNE, 1962) zu erklären, weisen die Autoren zurück.

Die Autoren selbst diskutieren eine Einschränkung der Allgemeinheit ihrer Befunde: Die Versuchspersonen nahmen beide Versuchsphasen als Teil desselben Experiments wahr. Auch THORNTON & JACOBS (1971) experimentierten mit gleichem Versuchsleiter und weiteren gleichen Cues in den beiden Phasen und verwendeten Schocks als aversive Reize. Als sie nach der Testphase ihre Versuchspersonen fragten, warum sie die Schocks ohne zu reagieren hinnähmen, antworteten die Versuchspersonen sinngemäß, sie könnten ja sowieso nichts machen, warum sie es denn dann versuchen sollten. MAIER & SELIGMAN (1976, S. 18) zitieren dies als kristallklare Evidenz für die reduzierte Motivation.

Zusammenfassend läßt sich zu den Ergebnissen der Humanversuche im triadischen Design sagen, daß der LH-Effekt - und vor allem das motivationale Defizit - mehrfach überzeugend repliziert wurde und mit Sicherheit kein Zufallsphänomen darstellt (z. T. greifen wir hier etwas vor). Der Transfer zwischen Problemlösungsaufgaben mit völlig unterschiedlichen Reizen und Operationen ist ebenfalls nachgewiesen und repliziert worden.

2.2.2.3 Erweiterung der LH-Theorie

2.2.2.3.1 Integration des Konfliktkonzepts

Ohne nähere Diskussion zeigt sich bei HIROTO & SELIGMAN (1975) eine Weiterentwicklung der LH-Theorie, die auch in den Überblicksarbeiten von SELIGMAN (1975) und MAIER & SELIGMAN (1974) nicht gewürdigt wird. Letztere definieren Unkontrollierbarkeit noch über die vom Versuchsleiter gesetzten Wahrscheinlichkeiten für denselben Verstärker:

$$K(R_i - C_j) = K(R_i - C_j); i = 1, 00 \text{ (Unkontrollierbarkeit)}$$

Im Experiment von HIROTO & SELIGMAN (1975) wurde das kognitive Hilflosigkeitstraining für die UK-Gruppe durch eine vorher festgelegte Reihenfolge von richtig-falsch-Antworten realisiert. Die Versuchsperson erhielt also für jede ihrer Antworten mit gleicher Wahrscheinlichkeit einen positiven Verstärker ('richtig') oder eine Strafe ('falsch'):

$$K(R_i - C^+) = K(R_i - C^-); i = 1, 00 \text{ (Unkontrollierbarkeit)}.$$

Nummehr ist Unkontrollierbarkeit definiert über die vom Versuchsleiter gesetzten gleichen Wahrscheinlichkeiten für gegenläufig wirkende Verstärker, die sich zudem in ihrer Wirkung aufheben. In Phase I wurde Hilflosigkeit definiert über die gleiche Wahrscheinlichkeit für denselben, deshalb gleichläufig wirkenden Verstärker für eine Reaktion und ihre Unterdrückung, mithin für inkompatible Operationen.

HIROTO & SELIGMAN (1975) sehen keinen Unterschied zu den ursprünglichen Operationalisierungen, weil Reaktion und Verstärker unabhängig voneinander sind (vgl. auch KLEIN et al., 1976).

Formal ähnelt die neue Form jedoch der Verhaltensgleichung für einen Approach-Avoidance-Konflikt: mit gleicher Wahrscheinlichkeit folgen auf eine Reaktion Belohnung und Bestrafung. Dies bedeutet, daß die Differenzierung in verschiedene nichtkontingente Treatments (vgl. z. B. KANFER & PHILLIPS, 1975, S. 287 und S. 368) in Phase II aufgegeben wird.

Wir wollen dies nachvollziehen (teilweise Replikationen für die o. a. Neuerung liegen von HIROTO, 1975; GATCHEL & PROCTOR, 1976, evtl. auch DWECK & REPUCCI, 1973, vor). Dies heißt von nun an, daß in jeder Situation, in der es zu einer kognitiven Repräsentation der Unabhängigkeit von Reaktion und Ergebnis kommt oder in der eine solche Kognition wirksam ist, ein LH-Effekt auftritt.

Die in den meisten LH-Überblicksarbeiten vorgenommene enge Koppelung zwischen Reaktionskontingenzraum (vgl. Abbildung 2, S. 57) und Hilflosigkeit ist deshalb irreführend. Die LH-Theorie wurde weiterentwickelt, und nunmehr reicht für das Eintreten eines LH-Effektes aus, wenn es durch Instruktion, allgemeine Kontrollorientierung oder tatsächliche entsprechende Vorerfahrung zur Ausbildung einer 'hilflosen' Kognition gekommen ist.

Damit ist der Begriff der 'kognitiven Repräsentation', wie wir ihn bislang verwendeten, zu modifi-

fizieren und weiterzuentwickeln. Unter Beibehaltung des Postulats der Analogie wollen wir die 'kognitive Repräsentation' mit den Elementen der Verhaltensgleichung beschreiben und eine Formel 'erfinden', für die die Bedingung der Reaktions-Ergebnis-Unabhängigkeit ebenfalls zutrifft oder auch nicht.

$$1. \quad S_i^D - R_1 \begin{matrix} K_i - C^- \\ K_i - C^+ \end{matrix}$$

$$2. \quad \begin{matrix} S_1^D - R_1 - K_i - C^- \\ S_2^D - R_1 - K_i - C^+ \end{matrix}$$

Die erste Formel entspricht den Verstärkungsbedingungen bei HIROTO & SELIGMAN, nur daß wir jetzt einen diskriminativen Reiz zusätzlich angenommen haben. Im zweiten Fall nehmen wir an, daß eine bestimmte Reaktion in der einen Situation bestraft, in der anderen belohnt wird: diejenigen, die zwischen den beiden Situationen diskriminieren, werden in der einen Situation die Reaktion R_1 letztlich unterdrücken; diejenigen jedoch, die die notwendige Diskrimination nicht leisteten, werden hilflos sein; für letztere trifft dann Formel 1 zu.

Damit sind wir auf eine Erklärungsvariante der LH-Theorie gestoßen, wie sie sich bereits bei der Diskussion der Tierexperimente andeutete. Wir wissen nicht, welche Funktion diskriminative Reize bei der Ausbildung des LH-Effektes haben und haben können. In unserem oben erfundenen Beispiel haben wir jedoch aufgezeigt, wie sich eine LH-Kognition durch Einführung diskriminierender Cues auflöst. Nun sagt die LH-Theorie zwar etwas aus über Interferenzen

beim Lernen von Reaktions-Ergebnis-Kontingenzen, wir finden aber keine Aussage zu Interferenzen beim Diskriminationslernen.

Formel 1 entspricht der Unkontrollierbarkeitsbedingung bei HIROTO & SELIGMAN (1975). Durch Einführung diskriminativer Reize könnten wir die Bedingung 'auflösen'. Im Umkehrschluß können wir folgern, daß in den LH-Experimenten zumindest bei den Versuchspersonen, bei denen ein LH-Effekt auftrat, keine 'Auflösung' stattfand.

WORTMAN & BREHM (1975) verweisen auf einige Experimente, in denen nach Wechsel von Versuchsleiter und Versuchsraum trotz Hilfllosigkeitstrainings kein LH-Effekt auftrat. Dort wirken dann Versuchsleiter und/oder Versuchsraum als diskriminativer Reiz.

Wenn wir weiter unten die LH-Experimente als spezielle Form von Problemlösungsexperimenten reinterpretieren, wird die unveränderte und nicht-differenzierte Situationskognition uns noch weiter beschäftigen.

Die Notwendigkeit, neben den Reaktions-Verstärkungs-Relationen auch die Situationskognition - die diskriminativen Reize des operanten Konditionierens - zu betrachten, macht man sich am besten am Beispiel des intermittierenden Verstärkens klar. Beim intermittierenden Verstärken folgt der Verstärker z. B. nur noch auf jede zweite Reaktion und wird sonst nicht eingesetzt. Betrachtet man eine Verstärkerquote von $K = .50$ und beschränkt man die Analyse

auf die Reaktions-Verstärkungs-Relationen, so entsteht eine Bedingung, in der 50 % der Reaktionen z. B. belohnt werden (C^+), aber zugleich $(100 - 50) \% = 50 \%$ der Reaktionen der Bedingung der Löschung unterworfen werden (ϕ^+).

So entsteht die folgende Formel:

$$K(R_i - C_j^+) = K(R_i - \phi_j^+) ; K = .50$$

Dies bedeutet, daß die Wahrscheinlichkeit für gegenläufig wirkende Verstärker gleich ist, also die (neue) Hilflosigkeitsbedingung erfüllt sein müßte. Über Hilflosigkeit beim intermittierenden Verstärken ist bislang nichts bekannt geworden, vielmehr scheinen die Organismen zu lernen, daß in der gegebenen Laborsituation eine Belohnung erfolgen kann, sofern vorher eine Reaktion erfolgte. Beim Übergang vom kontinuierlichen zum intermittierenden Verstärken bleibt diese Kognition erhalten, und es tritt somit kein LH-Effekt auf.

2.2.2.3.2 Gelernte Hilflosigkeit als Folge von Instruktionen

Bis hierher haben wir herausgearbeitet, daß für die Ausbildung eines LH-Effektes die Ausbildung der Kognition der Unabhängigkeit der Ergebnisse vom eigenen Handeln bei unveränderter Ausgangssituation ausreicht. Bislang haben wir diskutiert, wie die Ausbildung dieser Kognition durch Manipu-

lation der Vorerfahrung gefördert wird und welche Probleme dabei auftreten.

Wir stellen uns die Frage, ob entsprechende Vorerfahrungen nicht durch das für Humanexperimente spezifische Medium der Instruktion ersetzt werden kann.

Wir haben bereits erwähnt, daß in der Studie von HIROTO (1974; vgl. S. 23f.) eine 'chance'-Instruktion verglichen mit einer 'skill'-Instruktion die gleichen Effekte hatte wie Vorerfahrungen mit unkontrollierbaren aversiven Tönen verglichen mit 'kontrollierbaren' Vorerfahrungen.

THORNTON & POWELL (1974) berichten von mehreren Experimenten; in einem davon (Experiment II) wurden u. a. zwei Versuchsgruppen mit unkontrollierbaren Schocks in der Pretestphase konfrontiert. Eine dieser Gruppen wurde vor der Testphase instruiert, die Schocks seien von jetzt an kontingent mit ihren Reaktionen; die andere Gruppe blieb ohne neue Instruktion. In der Testphase zeigte sich nun, daß die 'Instruierten' sich signifikant von den 'Nicht-Instruierten' durch bessere Reaktionen unterschieden. Die 'Instruierten' unterschieden sich nicht mehr von der K-Gruppe. Die Autoren ziehen daraus den Schluß, daß Instruktion eine Analogie zur LH-Therapie von SELIGMAN, MAIER & GEER (1968) ist.

Andere Autoren fanden, daß unter 'skill'-Instruktionen die Versuchspersonen ihre Erwartungen hinsichtlich des weiteren Versuchsablaufs stärker der experimentellen Realität anpaßten als unter 'chance'-

Bedingungen (PHARES, 1957; vgl. auch MILLER & SELIGMAN, 1975; ROTTER & MULRY, 1965, KLEIN et al., 1976).

GLASS & SINGER (1972) berichten eine Serie von Untersuchungen, in denen allein die Instruktion, die Ergebnisse seien kontrollierbar, die Effekte der wirklichen Kontrollierbarkeit fast verdoppelten.

Besonders deutlich zeigt sich der Einfluß einer durch Instruktion induzierten Kontrollerwartung bei JAMES & ROTTER (1958) und bei ROTTER, LIVERANT & CROWNE (1961). JAMES & ROTTER zeigten, daß sich die vielbehauptete größere Extinktionsresistenz partiell (intermittierend) verstärkter Reaktionen nur unter 'chance'-Bedingungen zeigt; unter 'skill'-Bedingungen unterscheiden sich intermittierend und zu 100 % verstärkten Reaktionen nicht in ihrer Extinktionsresistenz. ROTTER, LIVERANT & CROWNE (1961) replizierten dieses Ergebnis, verwendeten jedoch keine 'skill'- und 'chance'-Instruktionen, sondern Aufgaben, deren Struktur im Sinne des kulturellen Hintergrundes der beteiligten Versuchspersonen 'chance'- oder 'skill'-artig waren. Die Erklärung dafür liegt darin (zusammenfassend LEFCOURT, 1966), daß unter 'chance'-Bedingungen ein Wechsel der Verstärkungsfrequenz von $K = 1.0$ zu $K = 0.0$ einem Wechsel in der Situation bzw. in den vom Versuchsleiter gesetzten Bedingungen signalisiert; für die intermittierend verstärkte Gruppe ist dies jedoch nicht sofort einsichtig, sie benötigt mehr Zeit, um den Situationswechsel zu erkennen. Unter 'skill'-Bedingungen jedoch, so vermuten die o. a. Autoren, attribuieren die Versuchspersonen ihre 'Mißerfolge' auf internale Faktoren wie Anstrengung oder Fähigkeit und benötigen daher beide die gleiche Zeit, um die veränderten Kontingenzen wahrzunehmen. Beide Gruppen tendieren daher zunächst weiter zu kontrollierendem Verhalten.

Zusammenfassend läßt sich sagen, daß das Unkontrollierbarkeitstreatment durch eine entsprechende 'chance'-Instruktion ersetzt werden kann. Gemeinsam

ist beiden Paradigmen, daß sich bei der Versuchsperson die Erwartung unkontrollierbarer Verstärker ausbildet (es zu einer entsprechenden kognitiven Repräsentation kommt).

Gemeinsam ist ihnen ferner, daß mit der Erwartung unkontrollierbarer Verstärker die Versuchspersonen weniger und weniger auf Veränderungen der objektiven Kontingenzen reagieren, sofern die Veränderung nicht überdeutlich ist.

2.2.2.3.3 Die Situationskognition in Experimenten zur Gelernten Hilflosigkeit und die Kausalattribution der Versuchspersonen

Eine deutliche Analogie zu den Konzepten der Kontrolle und des Locus-of-Control hat das Konzept der Kausalattribution. In dem Prozeß der retrospektiven Kausalattribution wird das eigene Handeln und das vorliegende Ergebnis durch Ursachenannahmen verknüpft.

WEINER et al. (1971) unterscheiden zwischen zwei Attributionsdimensionen, einer Stabilitätsdimension und einer Internalitätsdimension. Die durch die Dimensionen bestimmten Quadranten entsprechen vier unterscheidbaren Kausalfaktoren: Fähigkeit (internal/stabil), Anstrengung (internal/variabel), Aufgabenschwierigkeit (external/konstant) und Glück (external/variabel). Wir haben bereits herausgearbeitet, daß das Auftreten eines LH-Effektes an eine ebenso konstante wie 'unkontrollierbare' Situationskognition

gekoppelt ist. Obwohl die Ergebnisse von LH-Experimenten, in denen die Kausalattribution gemessen wurden, dem ersten Eindruck nach widersprüchlich erscheinen, scheint die unveränderte Kausalattribution ein gemeinsamer Nenner zu sein.

Zunächst aber wollen wir die Ergebnisse von LH-Experimenten darstellen, in denen die Kausalattribution als Variable verwendet wurde.

HIROTO (1974) hatte seinen Versuchspersonen in der 'chance'-Bedingung erklärt, die Kontrolle über die Ereignisse liege allein beim Versuchsleiter, aus der Sicht der Versuchsperson handele es sich um ein Ratespiel.

Dadurch war klar, daß ein zeitliches Zusammentreffen von Reaktion der Versuchsperson und Beenden des Tons external-variabel attribuiert wurde. Die 'chance'-Instruktion führte aber nicht nur zu einer (vermutlich) externalen und variablen Kausalattribution, sondern sie erwies sich zugleich als Äquivalent für Unkontrollierbarkeit.

DWECK & REPUCCI (1973) demonstrierten, daß 'hilflose' Kinder (die durch Lehrerurteil kategorisiert wurden) weniger Verantwortung für ihre Ergebnisse übernahmen als 'nicht-hilflose', also weniger internal attribuierten.

TENNEN & ELLER (1977) variierten ihre Versuchsbedingungen so, daß für die Versuchspersonen bestimmte Attributionsmöglichkeiten blockiert wurden. Ein LH-Effekt trat bei der Gruppe auf, die ein ausreichendes Hilflösigkeitstraining erhalten hatten und die Unabhängigkeit von Reaktion und Ergebnis internal attribuierten konnte. Eine andere Gruppe hatte die Möglichkeit, auf die Situation oder auf die Schwierigkeit der Aufgaben zu attribuiieren; bei ihr konnte kein LH-Effekt beobachtet werden. Leider haben die Autoren keinen Versuch unternommen, die Attribuition zu messen.

WORTMAN et al. (1976) manipulierten die Versuchsbedingungen ebenfalls so, daß die Wahrscheinlichkeit der Attribuition auf den einen Kausalfaktor wächst

und für den anderen sinkt. Sie maskierten ihr Experiment als Lärmforschung und ließen je zwei Versuchspersonen unter Lärm arbeiten. In einer Versuchsgruppe wurde Unkontrollierbarkeit für eine Versuchsperson realisiert, in einer anderen für beide. Sie postulierten, daß in ersterer Bedingung die Versuchspersonen internal attribuieren, in letzterer external attribuieren würden. Sie vermuteten, daß internale Attribuierung den LH-Effekt bedingt. Im Gegensatz zu den Erwartungen der Autoren schnitten die Versuchspersonen der ersten Bedingung in der Testphase jedoch besser ab - auch die 'hilflosen'. Leider versäumten auch diese Autoren, die Attribuierungen zu messen. Deshalb vermuten TENNEN & ELLER (1977), daß die Versuchspersonen der erstgenannten Bedingung auf den Lärm als konstanten Faktor attribuieren und es aus diesem Grunde zu keinem Transfer auf die Testphase (in der kein Lärm gegeben wurde) kam, während die Versuchspersonen der zweitgenannten Versuchsbedingung den Transfer vollzogen und darum hilflos reagierten. Eine weitere Erklärungsmöglichkeit ist, daß die Versuchspersonen der ersten Bedingung die Kompetenz des arwesenden Modells zum Anlaß nahmen, variabel zu attribuieren, um sich dann bei deutlich veränderten Versuchsbedingungen verstärkt anzustrengen. Eine Entscheidung darüber ist aber wegen der unterlassenen Messungen nicht möglich.

KLEIN et al. (1976) manipulierten die Attribuierungen in drei UK-Gruppen des triadischen Designs, indem sie den Versuchspersonen fiktive Vergleichsdaten über zu erwartendes Abschneiden bei Beginn der Pretestphase gaben. Der ersten UK-Gruppe wurde erzählt, daß Studenten in der Regel drei oder vier Problemlösungsaufgaben richtig lösen (Manipulation der internalen Mißerfolgsattribution), einer zweiten UK-Gruppe wurde in der Instruktion mitgeteilt, daß kaum ein Student auch nur eine Aufgabe löst (Manipulation der externalen Mißerfolgsattribution), für die dritte UK-Gruppe war keine Manipulation vorgesehen. Zusätzlich wurde bei allen Versuchspersonen die Depressivität mit dem BECK-Depressions-Inventary (BECK, 1967; BECK & BEAMES-DERFER, 1974) gemessen. Die Resultate der Messungen in der Testphase ergeben, daß die Attribuierungsmanipulation in den Verhaltensmaßen von Niedrig-Depressiven keinen Niederschlag fand, daß die durchweg schlechtere Leistung der Hoch-Depressiven durch Manipulation der externalen Attribuierung verbessert werden kann.

Die Interpretation der Autoren aller oben dargestellten Studien läßt eine konsistente Assoziation zwischen LH-Effekt und Attribuierungskomponenten kaum zu: Mal wird die externe Attribuierung (z. B. HIROTO, 1974), mal die interne Attribuierung (z. B. TENNEN & ELLER, 1977) zum LH-Effekt Beziehung gesetzt. Schaut man sich jedoch das tatsächliche Versuchsgeschehen - insbesondere die Instruktion - an, so kann man - wenngleich spekulativ - dennoch eine gewisse Konsistenz entdecken: In allen Fällen transferierten die Versuchspersonen die so generierten Erwartungen unverändert in die Testphase. Sicherlich spekulativ, aber nicht unplausibel, ist es daher, wenn wir annehmen, daß es in erster Linie darauf ankommt, daß die Kausalattribution unverändert ist, sei sie nun internal oder external (vgl. auch WORTMAN & BREHM, 1976, S. 231f.), stabil oder variabel.¹

¹ Nach Fertigstellung des Manuskripts und kurz vor dem Abgabetermin dieser Arbeit stießen wir auf drei Artikel, deren Autoren sich mit der Beziehung der Kausalattribution und der Gelernten Hilflosgigkeit - z. T. kritisch - auseinandersetzen: RIZLEY, R., Depression und Distortion in the Attribution of Causality, Journal of Abnormal Psychology, 1978, 87, 32-48; ABRAMSON, L.Y., SELIGMAN, M.E.P. & TEASDALE, J.D., Learned Helplessness in Humans: Critique and Reformulation, Journal of Abnormal Psychology, 1978, 87, 49-74; WORTMAN, C.B. & DINTZER, L., Is an Attributional Analysis of the Learned Helplessness Phenomenon Viable? A Critique of the ABRAMSON-SELIGMAN-TEASDALE Reformulation, Journal of Abnormal Psychology, 1978, 87, 75-90

Wenn wir annehmen, daß ein LH-Effekt nur dann auftritt, wenn die Kausalattribution unverändert bleibt, so argumentieren wir im Gleichklang mit unseren obigen Überlegungen zur unveränderten Situationskognition. Aus der Sicht beider Ansätze kann ein LH-Effekt dann vorhergesagt werden, wenn die Versuchsperson die Verstärker in der experimentellen Situation als unbeeinflußbar einschätzt und diese Einschätzung beibehält. Methodisch gesehen liegen zur Konstanz der Situationskognition keine empirischen Ergebnisse vor, wohl aber zur Konstanz der Attributionen.

2.2.2.4 Zusammenfassung und Diskussion

Seit Anfang der siebziger Jahre zeigt sich mit Beginn der Humanforschung eine Weiterentwicklung der LH-Theorie: Während die 'klassische' Version die Unkontrollierbarkeit eines Verstärkers dadurch definierte, daß er sowohl bei Realisation einer Reaktion als auch bei ihrer Unterdrückung auftritt, reicht es jetzt aus, wenn es irgendwie zur kognitiven Repräsentation der Unabhängigkeit von Reaktionen und Verstärkern kommt. Diese Unabhängigkeit kann durch eine allgemeine Kontrollorientierung im Sinne des Locus-of-Control gegeben sein, sie kann aber auch durch Instruktion oder Manipulation der Vorerfahrung induziert werden.

In Weiterführung des Konfliktparadigmas haben wir die Konstanz der Situationskognition als notwendige Bedingung für einen LH-Effekt eingeführt.

Wir wollen an dieser Stelle einen ersten Schritt zur Übertragung der LH-Theorie auf unser eigentliches Thema tun und die LH-Experimente als Problemlösungsprozeß aus der Sicht der Versuchsperson reinterpretieren.

Die Ausgangssituation im LH-Experiment ist klar: Der Versuchsleiter stellt eine Aufgabe (den Ton stoppen oder Konzepte finden). Gleichzeitig definiert er den Zielzustand (Beendigung des Tons oder die Endantwort 'richtig'). Es ist Aufgabe der Versuchsperson, Operationen zu finden, die Ausgangssituation und Zielzustand verbinden (z. B. Knopf drücken oder Exklusion irrelevanter Dimensionen). Unter 'unkontrollierbaren' Bedingungen wird dies vom Versuchsleiter systematisch verhindert, und die Versuchsperson lernt, daß es bei der gegebenen Ausgangssituation keine Operation gibt, die zum Zielzustand hinführt. In den meisten Experimenten wird dann bei gleichem Versuchsleiter und gleichem Versuchsraum eine neue Aufgabe gestellt (Hebel bewegen oder Anagramme lösen) und ein Zielzustand definiert (Abbrechen des Tons oder Aussprechen der Anagrammlösung). Die 'hilflosen' Versuchspersonen transferieren ihre Problemerkognition auf diese neue Aufgabe: Sie definieren den Zusammenhang der Merkmale der Ausgangssituation als unverändert und tun sich schwer, Ausgangssituation und Zielzustand zu verknüpfen. Als Folge davon sind die drei Aspekte des LH-Effektes beobachtbar: ein motivationales, kognitives und emotionales Defizit.

Damit wird klar, daß Gelernte Hilflosigkeit nichts anderes als eine Strategie ist.

Ein LH-Experiment können wir als Problemzustellung konzipieren, bei der Ausgangszustand und Zielzustand definiert sind und die verknüpfenden Operationen gesucht werden. In den Suchprozeß greift die Strategie der Gelernten Hilflosigkeit ein, und zwar entweder, indem der Suchprozeß abgeschwächt oder gar nicht erst initiiert wird - dies entspricht dem motivationalen Defizit -, oder indem der Suchprozeß auf irrelevante Operationen ausgedehnt wird (was zu verlängerten Latenzzeiten der Kriteriumsreaktion führt) oder indem sich die Beurteilung geeigneter Operationen verzögert - letztere beiden scheinen mehr dem kognitiven Defizit zu entsprechen.

Weder aus den Formulierungen der LH-Forscher noch aus den Beschreibungen der Experimente können wir Aussagen zu einer der drei genannten Möglichkeiten ableiten. Die LH-Theorie bleibt hier unspezifisch und global.

Wir haben deshalb den Eindruck, daß die LH-Theorie mehr ein System von Wenn-Dann-Formulierungen darstellt, das aus experimentellen Durchführungsrichtlinien besteht.

Man könnte sie nach Art eines Kochbuchs etwa so beschreiben:

Man nehme drei Gruppen von Versuchspersonen. Die eine wird kräftig frustriert. Nach einer kurzen

Pause lege man allen drei Gruppen Problemlöseaufgaben vor. In der frustrierten Gruppe zeigt sich dann ein Effekt, der LH-Effekt. Einen besonders ausgeprägten Effekt erhält man, wenn Versuchsleiter und Versuchsraum die gleichen bleiben.

Während in Phase I der Entwicklung der LH-Theorie sich die Forschung noch strikt an gegebene Richtlinien hielt, wird dies in Phase II aufgegeben. Dadurch war es möglich, das allgemeinere Konzept der Kontrollorientierung sensu ROTTER aufzugreifen.

Dies erwies sich zunächst als fruchtbar, führte jedoch zu einer 'Aufweichung' der 'klassischen' LH-Theorie und ging mit eher induktiven Forschungsansätzen einher.

Für unsere Absicht, die LH-Theorie auf das Gebiet der Prüfungsangst zu übertragen, hat dies zur Folge, daß wir uns weniger an der 'Theorie' orientieren und verstärkt darauf achten, was tatsächlich in den LH-Experimenten abläuft. Wir haben herausgearbeitet, daß es in der LH-Forschung weniger darauf ankommt, wie es zur kognitiven Repräsentation der Unabhängigkeit von Reaktion und Verstärker kommt, sondern mehr darauf, daß eine entsprechende Strategie wirksam ist. Wir haben herausgearbeitet, daß eine unveränderte Situationskognition und ein Mangel an diskriminativen Reizen im Zusammenhang mit dem LH-Effekt stehen muß. Wir haben herausgearbeitet, daß ein LH-Effekt auch als Folge einer geeigneten Instruktion auftreten kann, und auf ein Experiment verwiesen, in dem eine Instruktion dem LH-Effekt vorbeugte. Dies alles sind

Punkte, die die LH-Theorie der Phase II von der Theorie der Phase I unterscheiden.

In Phase III finden wir nun Bemühungen, das tatsächliche Geschehen exakter zu erfassen.

Das einzige vollständig durchgearbeitete Modell stammt von WORTMAN & BREHM (1975), die die Reaktanztheorie - oder wie IRLE (1975) schreibt: den 'Sachverhalt der Reaktanz' - mit der Theorie der Gelernten Hilflosigkeit integrieren.

2.2.3 Phase III: Humanversuche - Integration der Reaktanztheorie

2.2.3.1 Allgemeine Darstellung der Reaktanztheorie

Die Reaktanztheorie von BREHM (1966, 1972, Zusammenfassungen in IRLE, 1975, S. 372f., und WORTMAN & BREHM, 1975, S. 282f.) beinhaltet die Annahme, daß eine Person, deren Verhaltensspielraum, deren 'Freiheit' bedroht oder eingeengt ist, motiviert wird, diese Freiheit wieder herzustellen. Dieses Motiv nennt BREHM Reaktanz ('reactance').

Nach BREHM (1966) variiert die Reaktanz u. a. mit den Parametern:

- Kognition des Verhaltensspielraums (Reaktanz tritt dann auf, wenn ein Verhalten ver- oder behindert wird, für das die Versuchsperson Entscheidungsfreiheit beansprucht),
- Stärke der Bedrohung (eine Behinderung löst weniger Reaktanz aus als eine Verhinderung),

- Instrumentalität der Verhaltensweise (Behinderungen bei unwichtigen Reaktionen lösen weniger Reaktanz aus als Behinderungen bei wichtigen),
- Implikationen für andere, z. B. zukünftige, Verhaltensspielräume (je stärker die Folgen der Nichtausführung einer Reaktion sind, desto stärker fällt die Reaktanz aus).

Das Reaktanzmotiv wird (fast) nie direkt gemessen, sondern als intervenierende Variable über folgende Effekte erfaßt:

- Veränderung der Attraktivität des gefährdeten oder unerreichbaren Ergebnisses (die Reaktanztheorie sagt - im Gegensatz zur Dissonanztheorie - eine Aufwertung der eliminierten Alternative vorher),
- direkte Ausföhrung der bedrohten Verhaltensalternative (vgl. hierzu die Unterbrechungstheorie von MANDLER & WATSON, 1966, nach der eine mögliche Reaktion auf die Unterbrechung einer Verhaltenssequenz die Wiederholung oder Fortsetzung der Sequenz ist),
- Realisation von Ersatzresponses (vgl. MANDLER & WATSON, 1966),
- feindliche und aggressive Geföhle gegenüber dem Verantwortlichen.

Die Analogie zwischen LH-Theorie und Reaktanztheorie dürfte auf der Hand liegen. In beiden Fällen wird der Versuchsperson die Kontrolle über ein Ergebnis entzogen, beide definieren kognitive Konstrukte, die über ihre Auswirkungen analysiert werden. Beide Theorien lassen Fragestellungen offen, beschreiben jedoch mehrfach replizierte Effekte.

Während die LH-Theorie ein Absinken der Verhaltensfrequenz als Folge des Kontrollverlustes vorher-
sagt, folgt aus der Reaktanztheorie die Prognose
eines Anstiegs. Die LH-Theorie sagt passives Ver-
halten und depressive Stimmungen vorher, die Reak-
tanztheorie aktives Bemühen und Aggressionen.

2.2.3.2 Darstellung des integrativen Modells

WORTMAN & BREHM (1975) legen nun ein integratives
Modell vor und heben die scheinbaren Widersprüche
auf:

"Wenn eine Person erwartet, Ergebnisse, die ihr
relativ wichtig sind, beeinflussen oder kontrol-
lieren zu können, diese sich aber als unkontrol-
lierbar erweisen, dann wird die psychologische
Reaktanz angeregt. Auf Individuen, die anfangs
Kontrolle erwarten, sollten die ersten Durchgänge
eines Hilflosigkeitstrainings deshalb als Bedrohung
ihrer Freiheit wirken. Sie sollten eine erhöhte
Motivation, Kontrolle auszuüben, erfahren, und
die Performanz sollte sich verbessern. Die Reak-
tanz sollte um so stärker ausfallen, je wichtiger
die unkontrollierbaren Ergebnisse für die Person
sind. Trotz seiner erhöhten Motivation lernt das
Individuum während des andauernden Hilflosigkeits-
trainings jedoch allmählich, daß es das Ergebnis
nicht kontrollieren kann. Wenn eine Person schließ-
lich überzeugt ist, daß sie die Ergebnisse nicht
kontrollieren kann, stellt sie die Bemühungen ein.
Reaktanz geht der Hilflosigkeit bei Individuen
voraus, die anfangs Kontrolle erwarten. Je stärker
diese Kontrollerwartung, desto länger dauert es
vernünftig, bis eine Person hilflos wird, und
desto höher wird die Frequenz kontrollierender
Reaktionen und die Ausdauer dabei sein... Bei In-
dividuen, die erst gar keine Kontrolle erwarten,
sollte keine Reaktanz angeregt werden, und sie
sollten sich relativ schnell hilflos verhalten,
wenn sie mit unkontrollierbaren Verstärkern kon-
frontiert werden." (WORTMAN & BREHM, 1975, S. 308,
übersetzt vom Verfasser)

Im Hilflosigkeitstraining reagieren die Versuchspersonen zunächst auf den Kontrollverlust mit aktiven Versuchen, die Kontrolle wiederherzustellen, erst zu einem späteren Zeitpunkt lernen sie die Unabhängigkeit von Reaktion und Ergebnis und werden passiv. Dieser Zeitpunkt verzögert sich um so mehr, je größer die Kontrollerwartung am Anfang ist. Ferner nehmen WORTMAN & BREHM an, daß sowohl der Reaktanz- als auch der LH-Effekt desto stärker ausfallen, je 'wichtiger' das Ergebnis für die Versuchsperson ist.

In Abbildung 4 sind die Grundannahmen der Autoren illustrativ zusammengefaßt.

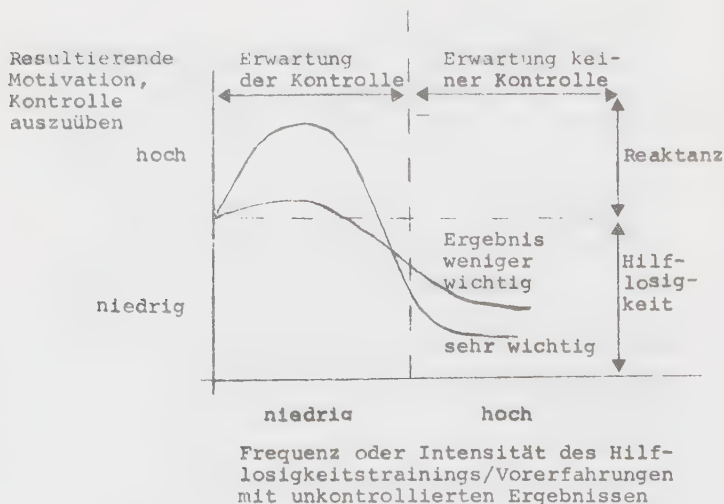


Abbildung 4: Das integrative Modell von WORTMAN & BREHM (angelehnt an WORTMAN & BREHM, 1975, S. 309)

Das integrative Modell wird durch einzelne experimentelle Daten gestützt. WORTMAN & BREHM diskutieren die entsprechenden Tierstudien, wir gehen darum nicht näher darauf ein.

WORTMAN & BREHM zitieren drei Humanexperimente, die ihr Modell stützen (SHABAN & WELLING, zitiert nach GLASS & SINGER, 1972, S. 122-130; KRANTZ et al., 1974; ROTH & KUBAL, 1976).

Bereits bei HIROTO & SELIGMAN (1975) zeigte sich, daß unter bestimmten Bedingungen erst eine Erhöhung des Hilflosigkeitstrainings den LH-Effekt hervorbringt.

THORNTON & JACOBS (1972) und ROTH & BOOTZIN (1974) fanden in ihren UK-Gruppen statt eines Leistungsdekrements eine Leistungssteigerung.

ROTH & KUBAL (1976) experimentierten mit sieben Gruppen. In der Pretestphase lösten die Versuchspersonen Konzeptbildungsaufgaben (mit Ausnahme der N-Gruppe, die sofort mit der Testphase begann). Von den übrigen sechs Gruppen erhielten zwei Gruppen kontingente Antworten (= KG), vier Gruppen erhielten nicht-kontingentes Feedback (= UKG). Von letzteren wurden je zwei einem Trainingsdurchgang (= UKG1), die beiden anderen je drei Durchgängen unterworfen (= UKG3). Die Wichtigkeit wurde durch die Instruktion manipuliert: Der einen Hälfte wurde erzählt, daß die Trainingsaufgaben den akademischen Erfolg prognostizierten (= W), der anderen Hälfte wurde ein Spiel angekündigt (= w). Die sieben Grup-

pen des Designs waren also: KG-w, KG.W, UKG1-w, UKG1-W, UKG3-w, UKG3-W und N-Gruppen.

Die Testphase wurde als völlig unabhängiges Experiment angekündigt und auch von einem anderen Versuchsleiter in einem anderen Raum durchgeführt. Der Versuchsleiter der Testphase war über die Zuordnung zu den Gruppen der Pretestphase nicht informiert.

Für die Verhaltensmaße fanden ROTH & KUBAL Ergebnisse, die im Einklang mit dem integrativen Modell stehen: die UK3-Gruppen reagierten hilflos, die UK1-Gruppen reaktant. Die Manipulation der Wichtigkeit war nicht erfolgreich.

Interessant an diesem Experiment ist auch die Analyse des umfangreichen postexperimentellen Fragebogens. Danach unterschieden sich die Gruppen nicht u. a. in der Einschätzung der Beziehung zwischen den beiden Testphasen, in der Ängstlichkeit, in der Ermüdung, in ihrer Anstrengung und Engagiertheit. Die UKG3-w-Versuchspersonen fühlten sich hilfloser, die UKG3-W schätzten die Probleme eher unlösbar ein und ärgerten sich mehr. Die Versuchspersonen der UKG1-W empfanden gegenüber dem Versuchsleiter der Testphase signifikant positivere Gefühle.

TENNEN & ELLER (1977) replizierten das eben erwähnte Experiment teilweise. Sie verwendeten neben N-Gruppen, K-Gruppen und UK1-Gruppen zwei UK3-Gruppen, insgesamt also fünf Gruppen. Den

Versuchspersonen wurden zwei verschiedene aufeinanderfolgende Experimente angekündigt, die auch in verschiedenen Räumen durchgeführt wurden (von verschiedenen Versuchsleitern in Pretest- und Testphase schreiben TENNEN & ELLER nichts). Die Ergebnisse waren in Übereinstimmung mit dem integrativen Modell. Die Skalierung von Stimmungen vor und nach der Pretestphase zeigte, daß alle Gruppen weniger nervös wurden, die K-Gruppe sich signifikant weniger traurig fühlte und die UK1-Gruppe sich signifikant mehr ärgerte.

Insgesamt haben wir den Eindruck, daß nichts gegen das integrative Modell von WORTMAN & BREHM (1975) spricht, wenngleich auch erst wenig empirische Daten vorliegen. Ähnliche Beobachtungen, daß mit dem Entzug der Kontrolle über ein Ereignis die Verhaltensfrequenz zunächst steigt, um erst nach einiger Zeit abzusinken, sind ja auch aus dem Bereich der Löschung operanten Verhaltens - einem Spezialfall an Hilflosigkeit - bekannt (KANFER & PHILLIPS, 1975, S. 322).

Einhergehend mit der Integration beider Theorien stoßen wir auf das Konzept der 'Wichtigkeit' des Verstärkers. Auf einem sehr allgemeinen Niveau ist plausibel, daß der funktionale Stellenwert des von der Versuchsperson zu kontrollierenden Verstärkers nicht gleichgültig ist. So ist es nach Ansicht des Verfassers ein Unterschied, ob der nächste Angehörige stirbt (Kontrollverlust über soziale Verstärker) oder ein penetranter Piepton nicht abgeschaltet werden kann.

Das Verdienst von WORTMAN & BREHM ist sicherlich, die Theoriebildung auf die Pretestphase ausgedehnt zu haben und damit differenziertere Aussagen zu ermöglichen. Damit erhöhen sie die interne Validität (CAMPBELL & STANLEY, 1963, deutsch: 1970; vgl. MERTENS, 1975, S. 23f.).

Für unser eigentliches Vorhaben, der Übertragung der LH-Theorie auf das Gebiet der Prüfungsangst, ist jedoch die externe Validität - das Ausmaß der Übertragbarkeit der experimentellen Befunde auf real-life-Situationen interessanter. Und eine Theorie, die den Anspruch erhebt, ein Erklärungsmodell für Depressionen zu liefern (SELIGMAN, 1975; SELIGMAN, 1974; MILLER & SELIGMAN, 1973; SELIGMAN, 1973), muß sich natürlich dieser Frage stellen.

Ein Leser, der den Darstellungen des Verfassers bis hierhin gefolgt ist, wird sich nun vielleicht fragen, was - bei aller Kritik der Theorie - denn nun von ihr zusammenfassend zu halten sei. Darauf eine Antwort in der üblichen 'ausgewogenen' Form geben zu wollen, halten wir für verfehlt. Die Theorie der Gelernten Hilflosigkeit befindet sich in einer sehr schnellen Entwicklung. Der Beginn der LH-Forschung mit sehr genau kontrollierten Tierversuchen brachte Ergebnisse, die nicht wegdiskutiert werden können: Sobald ein Organismus lernt, daß seine Reaktionen und bestimmte Verstärker in keinem Zusammenhang stehen, hat dies Folgen, die man als motivationales, kognitives und emotionales Defizit beschreiben kann. Beginnend mit Phase II

wird die Theorie einerseits um Komponenten erweitert, andererseits wird sie unpräziser. Am deutlichsten wird dies in Phase III beim integrativen Modell; zwar sind jetzt Hypothesen hinsichtlich des gesamten Versuchsablaufs ableitbar, und somit ist die Möglichkeit gegeben, das gesamte Versuchsgeschehen vollständig zu erfassen, andererseits bringt gerade das integrative Modell eine gewisse Beliebigkeit der Post-hoc-Interpretation mit sich: ganz egal, welche Ergebnisse wir in der Testphase erhalten, wir können sie 'erklären'; steigt die Verhaltensfrequenz, interpretieren wir dies als Reaktanzeffekt, sinkt sie, so sprechen wir von LH-Effekt.

Für die weitere Entwicklung der Theorie wird es unerlässlich sein, sie zu präzisieren. Gelingt dies nicht, dann wird 'Gelernte Hilflosigkeit' schon bald zu einem vagen, mehrdeutigen Schlagwort werden. Zum gegenwärtigen Zeitpunkt empfiehlt es sich, sehr sorgfältig und kritisch mit dieser Theorie zu arbeiten.

Festzuhalten ist, daß unter bestimmten Umständen ein LH-Effekt auftreten kann, wenngleich mit den bisher dargestellten Ansätzen die Bestimmung der bestimmten Umstände nur global möglich ist.

2.2.4 Die äußere Validität der Experimente zur Gelernten Hilflosigkeit

Der Aspekt der äußeren Validität (CAMPBELL & STANLEY, 1963) betrifft die Frage, inwieweit die experimentell gewonnen Befunde auf Alltagssituationen oder vorgegebene klinische Probleme übertragen werden können.

Für die Übertragbarkeit des LH-Modells auf das Gebiet der Depression nennt SELIGMAN vier Prüfkategorien: "There are four relevant lines of evidence for asserting that two phenomena are similar:

- 1) behavioral and physiological symptoms,
 - 2) etiology or cause, 3) cure, and 4) prevention."
- (SELIGMAN, 1975, S. 80)

Wir hatten folgende Kriterien für die Übertragbarkeit des LH-Modells auf den Bereich der Prüfungsangst genannt:

- gleiche motorische Symptome (motorisch-verhaltensmäßige Betrachtungsebene),
- gleiche physiologische und affektive Symptome (organisch-physiologische und subjektive Betrachtungsebenen),
- gleiche Ätiologie (historischer Aspekt),
- gleiche systemfunktionalistische Parameter (symbolisch-kommunikative Betrachtungsebenen).

Die von SELIGMAN aufgezählten Prüfkategorien der Therapie und der Prävention wollen wir aus dieser Arbeit ausklammern. Ziel dieser Arbeit ist die Ent-

wicklung eines adäquaten Modells der Prüfungsangst (wenngleich ursprünglich mit dem Ziel der Prävention, Prophylaxe). Die Therapie der Prüfungsangst würde ein eigenständiges Kapitel mit besonders kritischer Diskussion erfordern, die zu unserem Thema - gemessen an dem notwendigen Umfang - sehr wenig beiträgt.

Den Nachweis gleicher affektiver Symptome - sei das LH-Modell nun auf depressives Verhalten oder auf Prüfungsangst zu übertragen - halten wir für unerlässlich.

Die affektiven (Fragebogen-) Symptome des LH-Experiments sind allenfalls ansatzweise und unsystematisch untersucht (z. B. ROTH & KUBAL, 1976; TENNEN & ELLER, 1977). Bislang ist uns kein Experiment bekannt, das die Veränderung von Depressionsscores durch nicht-kontingentes Treatment untersucht. Die (deshalb auch so ausführlich dargestellten) Versuchsbeschreibungen lassen Raum zu der Annahme, daß die affektiven Reaktionen im LH-Experiment allenfalls die Wucht der Samstagabendenttäuschung des Lottotippers erreichen.

Die symbolisch-kommunikativen Analogien von LH-Effekt und depressivem Verhalten bleiben völlig undiskutiert: die Instrumentalität der gesamten Laborsituation bleibt offen. Die sozialen Konsequenzen einer depressiven Entwicklung erreichen z. T. existentielles Niveau: sozialer Rückzug bis zur völligen Isolation und zum Selbstmord, generelles Absinken der Leistungs- und Belastungs-

fähigkeit und damit einhergehend Schwierigkeiten im Arbeitsleben, Verärgerung der unmittelbaren Umgebung durch Skrupulosität, Entschlußlosigkeit und Egozentrismus etc. Selbst wenn man konzidiert, daß ein Experiment mit solchen Folgen ethisch nicht legitimiert werden kann (vgl. WORTMAN & BREHM, 1975, S. 326), muß man feststellen, daß im Humanversuch der LH-Effekt nicht einmal das Ausmaß leichteren Liebeskummers erreicht. Es erscheint uns fast überflüssig anzumerken, daß die potentiellen Risiken für die Versuchspersonen nicht anekdotisch erwähnt oder gar explizit diskutiert werden.

Folgerungen, die sich aus der Interpretation des Experiments als sozialer Situation ergeben, dessen notwendiger Bestandteil eine Interaktion zwischen Versuchsleiter und Versuchsperson ist, werden unnötig ausgeklammert. Die Frage der äußeren Validität, nämlich die Übertragbarkeit der Situationsstruktur des Experiments auf Alltagssituationen, wird nicht thematisiert. MERTENS (1975) nennt eine Fülle möglicher Ansätze, die hier diskutiert werden könnten.

Nun hat die LH-Forschung nachgewiesen, daß z. T. unterschiedliche Reaktions-Verstärkungs-Paradigmen zum LH-Effekt führen. Wir meinen, daß alle methodischen Varianten, die zur Ausbildung eines LH-Effektes führen können, unter eine gemeinsame Meta-Theorie subsumiert werden können: die Theorie der Kommunikation zwischen Versuchsleiter und Versuchsperson im LH-Experiment. Wir meinen ferner, daß von daher die Interaktion in Prüfungen und das Auftreten von Prüfungsbelastungen erklärt werden kann.

2.3 Die Übertragung der Theorie der Gelernten Hilflosigkeit auf die Prüfungsinteraktion

Dies nämlich ist das Gemeinsame von LH-Experiment und mündlicher Prüfung: aus der Sicht der Versuchsperson und des Prüflings handelt es sich um eine Problemstellung, bei der Ausgangssituation und Zielzustand definiert sind und die verknüpfenden Operationen zu suchen sind. Auf bislang ungeklärte Weise greifen Strategien in den Suchprozeß ein. Die LH-Theorie besagt, daß 'hilflose' Strategien durch entsprechende Vorerfahrungen und durch entsprechende Kontrollorientierungen sich ausbilden. 'Hilflose' Strategien sind gekoppelt an ein motivationales Defizit - i. e. passives Verhalten -, ein kognitives Defizit - i. e. die Unfähigkeit, aus Rückmeldungen (Verstärkern) zu lernen- und an ein emotionales Defizit - das noch zu besprechen ist.

Nun ist das Problem des Prüflings in erster Linie ein kommunikatives: es gilt, in der Kommunikationsstruktur der Prüfung ein - bislang nicht näher zu definierendes - Kommunikationsverhalten zu realisieren. Dies lenkt unsere Aufmerksamkeit auf die Kommunikationstheorie von WATZLAWICK, BEAVIN & JACKSON (1969). Dabei finden wir, daß im LH-Experiment und in einer Prüfung die gleichen Kommunikationsstrukturen zu finden sind.

In mündlichen Diskussionen unseres Themas erfolgte an dieser Stelle häufig der Einwand, in Prüfungen gehe es doch in erster Linie um das Wissen des Kandidaten, genau dies werde geprüft. In dieser Absolutheit und mit dieser Priorität - 'genau dies' und 'in erster Linie' - ist dieser Einwand falsch. In

einer mündlichen Prüfung kann sich ein Prüfer allein an dem orientieren, was der Kandidat sagt. Wenn wir jemanden, der alles weiß, als 'Genie' bezeichnen, dann wird dieses Genie dennoch 'durchfallen', wenn es sein Alles-Wissen nicht in der Prüfungskommunikation demonstriert. Eine logische Konsequenz der Priorität der Kommunikationskompetenz ist, daß ein Genie, das in seiner All-Überlegenheit aus der Sicht des Prüfers unverständlich redet, ebenfalls durchfällt.

2.3.1 Symbolisch-kommunikative Betrachtungsebene

2.3.1.1 Eine kommunikationstheoretische Reanalyse der Experimente zur Gelernten Hilflosigkeit

Wir wollen versuchen, eine kommunikationstheoretische Betrachtungsweise des experimentellen Geschehens auf das bereits ausführlich beschriebene Experiment von HIROTO & SELIGMAN (1975) exemplarisch zu übertragen (vgl. hierzu MERTENS, 1975, S. 152f.; vgl. WATZLAWICK, BEAVIN & JACKSON, 1969).

MERTENS (1975) analysiert die Funktion der Versuchsperson im Experiment u. a. aus einem Blickwinkel, den wir im sozialwissenschaftlichen Einstieg in unser Thema als systemfunktionalistisch bezeichnet haben. Er geht davon aus, daß Versuchsleiter und Versuchsperson im Experiment so kommunizieren, daß z. B. der Versuchsleiter die Verstärker, die Versuchsperson die Reaktionen liefert. Er sieht die Versuchsperson in der Rolle des Ohnmächtig-Ausgelieferten:

"So signalisiert die in unpersönlichem Ton vorge-tragen Instruktion von vornherein, daß der V1 an

keinem näheren Kontakt mit der Vp interessiert ist, daß er das Recht hat, die Ausführung bestimmter unsinniger, langweiliger oder grausamer Aktivitäten zu fordern, daß ein Widerspruch oder eine Auseinandersetzung über den Inhalt der Kommunikation unerwünscht und unerlaubt ist usf. Aufgrund des methodologischen Regulativs der Standardisierung ist eine explizite Thematisierung und Problematisierung des Beziehungsaspekts unzulässig. Geschieht dies dennoch, ..., so werden solche Fälle aus der statistischen Analyse ausgeschlossen... Die Vp ... kann sich lediglich als gefügige Vp definieren, die bereitwillig die Instruktionen des Vls ausführt. Weigert sie sich offen, die Kompetenz des Vls anzuerkennen, so wird das Experiment abgebrochen... Wenngleich es richtig ist, daß unter dem Aspekt der sozialen Macht dem Vl potentiell mehr Möglichkeiten der Beziehungsdefinition zur Verfügung stehen, so ist es ihm wegen des Zwangs zur Standardisierung nicht erlaubt, Rückfragen über die Akzeptierung seiner Selbstdefinition einzuholen. So gesehen, kann die implizite Metakommunikation - die Beziehungsdefinition - niemals in der Interaktion von Vl und Vp expliziert werden." (MERTENS, 1975, S. 153f.)

Während also der Versuchsleiter in der Instruktion inhaltlich kommuniziert, definiert er gleichzeitig eine schiefe Rollenverteilung über die Beziehungskommunikation. Zusätzlich verbietet der Versuchsleiter der Versuchsperson die Metakommunikation, d. h. die Versuchsperson darf den Beziehungsaspekt nicht problematisieren. Dem Versuchsleiter selbst ist die Metakommunikation (zumindest während des Experiments) freilich auch verboten.

Diese Restriktion auf eine asymmetrische, komplementäre Interaktion, in der der eine Partner sein Verhalten primär nach den Bedürfnissen des anderen ausrichtet, keine eigenständigen Handlungsinter-

tionen verfolgen kann und über die Beziehung auch nicht metakommunizieren kann, ähnelt, wie MERTENS schreibt, den Interaktionen in totalen Institutionen.

Wir wollen den soeben dargestellten Gedanken auf das Experiment von HIROTO & SELIGMAN (1975) übertragen. In dem Experiment wurde der Transfer des LH-Effektes von der kognitiven/instrumentellen Pretestphase auf eine kognitive/instrumentelle Testphase untersucht (vgl. S. 84f.)

Der Gruppe wurde kommuniziert, sie könnten einen Ton vermeiden bzw. ein Konzept bilden. Das gleiche geschah im 'unkontrollierbaren' Treatment. Bei der instrumentellen Aufgabenstellung ('to stop the noise') wurde explizit verboten: "Taking the ear-phones off or dismantling the apparatus is not the way to stop the noise." (HIROTO & SELIGMAN, 1975, S. 317). Kein einziger Satz der 'instrumentellen' Instruktion enthält ein Pronomen, das sich auf die Person des Versuchsleiters bezieht. Bei der kognitiven Aufgabenstellung ('to figure out what the answer is') wurde den Versuchspersonen kommuniziert, daß der Versuchsleiter ('I') die Lösungen willkürlich gewählt hätte und die Wahlen der Versuchspersonen mit 'richtig' oder 'falsch' bewerten würde. Ihre Aufgabe sei es, den richtigen Wert herauszufinden.

Ohne Zweifel ist die Instruktion unpersönlich; der Versuchsleiter definiert eine komplementäre Interaktion; der Versuchsperson bleibt nichts anderes

übrig, als die Beziehungsdefinition zu akzeptieren. Dies kann von der Versuchsperson nur durch eine einzige Reaktion unterlaufen werden: sie geht aus dem Feld, d. h. sie bricht das Experiment ab.

Für die Analyse der Situation der UK-Gruppe müssen wir nun ein kommunikationstheoretisches Axiom berücksichtigen: das Axiom der Interpunktion. Es besagt, daß der Kommunikationsablauf nach Anfang und Ende, Reiz und Reaktion strukturiert wird und daß dies die Beziehungen zwischen den Interaktionspartnern bestimmt. Im Lernexperiment liefert der Versuchsleiter 'Reiz' und 'Verstärkung', die Versuchsperson liefert die 'Reaktionen' (vgl. MERTENS, 1975, S. 155).

Für die N-Gruppe und die K-Gruppe stellt sich der Versuchsablauf nun so dar, daß die anfängliche metasprachliche Beziehungsdefinition durch den Versuchsablauf und die damit einhergehende Interpunktion der Beziehung bestätigt wird. Um es deutlich zu formulieren: die Beziehung der Interaktionspartner wird nicht nur metasprachlich, sondern auch durch die Struktur des Kommunikationsablaufes bestimmt. Akzeptieren die Versuchspersonen von N- und K-Gruppe die Beziehungsdefinition, so bleibt die Beziehung während des Versuchs ungestört.

Die Interaktionsstruktur für die UK-Gruppe sieht anders aus. Bei ihr werden die durch die Instruktion kommunizierten Inhalte durch den Versuchsablauf nicht 'bestätigt'. Prinzipiell liegt daher die Möglichkeit eines Beziehungskonflikts vor, einer Störung der Interaktion.

So gesehen ergibt sich, daß sich die UK-Gruppe von den beiden anderen Gruppen des triadischen Designs insofern unterscheidet, als bei ihr eine Kommunikationsstörung wirksam ist. N- und K-Gruppe unterscheiden sich in dieser Hinsicht nicht. Die Kommunikationsstörung besteht in einer gestörten Beziehung zwischen den Partnern des Systems 'Experiment' (die allerdings in den meisten LH-Experimenten vom Versuchsleiter initiiert wird und beabsichtigt ist).

Wir können demnach sagen, daß es im LH-Experiment nicht nur um bestimmte Problemlösungen geht, die bei der UK-Gruppe vom Versuchsleiter systematisch vereitelt werden, sondern wir können auch definierte Kommunikationsmuster feststellen. Wie bei jedem Experiment handelt es sich um eine zeitlich limitierte Interaktion, bei der der Versuchsleiter eine komplementäre Beziehung definiert, bei der Metakommunikationen verboten sind, bei der es verboten ist, aus dem Felde zu gehen, bei der eine Kommunikationsverweigerung unmöglich ist und bei der der Versuchsleiter Reiz und Verstärkung liefert, die Versuchsperson die dazugehörigen Reaktionen. Bei der UK-Gruppe wird es der Versuchsperson aber unmöglich gemacht, Reiz und Reaktion durch geeignete Reaktionen zu verknüpfen. Die UK-Gruppe findet sich in der Situation einer Doppelbindung: sie wurde zunächst veranlaßt, eine bestimmte Kommunikationsstruktur zu akzeptieren, dann aufgefordert, ein Problem zu lösen, was sich im zeitlichen Verlauf aber als 'unlösbar' herausstellt. In einem neuen experimentellen Teil wird dann ein LH-Effekt

bei der UK-Gruppe unter zwei Bedingungen auftreten: die Kommunikationsstruktur wird weiterhin nicht in Frage gestellt, und die Reiz-Verstärkungs-Relation wird als nicht-existent kogniziert.

Nun hat die LH-Forschung aber gezeigt, daß z. B. eine externale Kontrollorientierung ebenfalls zum LH-Effekt führt. Hier eben zeigt sich der Vorteil einer systemtheoretischen - in dieser Arbeit kommunikationstheoretischen - Betrachtung. Die paradoxe Handlungsaufforderung in der Doppelbindung muß nicht notwendig durch einen Versuchsleiter induziert werden - dies ist nur eine mögliche Variante -, sie kann ebenso von der Versuchsperson in das Kommunikationssystem Experiment eingebracht werden.

2.3.1.2 Ableitung der Hypothesen zum Verhältnis von Vorerfahrung, Prüfungsstrategie und Leistungsdaten

Eine Versuchsperson, die von vornherein nicht erwartet, Kontrolle über Ergebnisse realisieren zu können, schafft zunächst einmal die gleichen Bedingungen, wie sie nach einem Hilflosigkeitstraining herrschen: die Kommunikationsstruktur wird nicht in Frage gestellt, und die Reiz-Verstärkungs-Relation wird als nicht-existent kogniziert. Dies ist aber auch schon ein Teil des LH-Effektes.

Es ist an der Zeit, jetzt wieder zum Thema dieser Arbeit zurückzukehren.

Ein Prüfungskandidat befindet sich nämlich in der

gleichen Lage wie eine Versuchsperson im LH-Experiment.

Ebenso wie eine Versuchsperson sich zur Teilnahme am LH-Experiment bereit finden kann (oder auch nicht), ebenso kann sich ein Prüfling zur Prüfung melden (oder auch nicht). In dem Augenblick, in dem die Teilnahme 'besiegelt' ist - bei der Versuchsperson z. B. durch ein mündliches 'okay', beim Prüfling durch Einreichen der Unterlagen - muß er sich bestimmten Regeln unterwerfen: die Versuchsperson muß die asymmetrische Beziehung mit Metakommunikationsverbot akzeptieren, der Prüfling muß plötzlich - z. T. 'willkürliche' - Bestimmungen beachten und sich letztlich den Regeln der mündlichen Prüfung unterwerfen. Das Prüfungsstereotyp haben wir im ersten Kapitel beschrieben: eine übermächtige Institution wird durch den Prüfer repräsentiert, die Prüfung verläuft 'standardisiert' und ritualisiert, die Ausgangssituation zu Beginn der Prüfung ist ebenso definiert wie der Zielzustand, der sich nach der Prüfung ergibt. Aufgabe des Prüflings ist es, für das Kommunikationssystem Prüfung Operationen zu finden, die Ausgangssituation und Zielzustand zu verbinden. Wir wollen uns die Arbeit erleichtern und den Zielzustand mit der Zensur operationalisieren.

In der mündlichen Prüfung liefert der Prüfer die Themen (Fragestellungen) und die Rückmeldungen, der Kandidat die Stellungnahmen (Antworten). Bei der Analyse der Situation der Prüfung haben wir bereits aufgezeigt, daß Themen und Rückmeldungen

durch den Prüfer nicht unabhängig von den Stellungnahmen des Kandidaten sind. Die Abhängigkeit der Rückmeldungen des Prüfers (z. B.: "ja, richtig!" oder "nein, nicht ganz!") von den Stellungnahmen des Kandidaten ist dabei unmittelbar plausibel. Aber auch die Auswahl der Themen ist von den Prüflingsbeiträgen zum Kommunikationsprozeß abhängig, sei es nun, daß z. B. der Prüfer bei exzellenten Antworten schwierigere Fragen stellt, um eine gute Zensur begründen zu können, sei es, daß er z. B. auf Reizworte reagiert ("Ach, Sie sagten da gerade... Können Sie das erläutern?"). Zielzustand des Prüfers ist dabei eine ausreichende Menge von Indizien für eine Zensur, und, um diese zu erhalten, steht die Steuerungsfunktion (eine definierte Menge von Operationen) zur Verfügung. Der Zielzustand des Kandidaten unterscheidet sich dabei insofern von dem des Prüfers, als er eine bestimmte Zensur - eine möglichst gute - erreichen will. Dies kann der Kandidat dadurch realisieren, daß er in den Steuerungsprozeß eingreift, daß er dazu übergeht, von sich aus Indizien anzubieten.

Daraus folgt, daß er aktiv kontrollierendes Verhalten realisieren muß.

Wir können daraus die Hypothese ableiten, daß erfolgreiche Kandidaten aktive und kontrollierende Strategien realisieren, während nicht-erfolgreiche passiv und hilflos agieren (Hypothese 4). Dies steht in Übereinstimmung mit den bisherigen empirischen Ergebnissen, wonach schlechte Zensuren mit passiver Hilflosigkeit assoziiert sind, mit

dem Gefühl, in der Prüfung ohnmächtig ausgeliefert zu sein, mit dem Gefühl, der Prüfer wolle den Kandidaten fertigmachen (vgl. S. 40f.)

Bei der Diskussion und Darstellung der sozialwissenschaftlichen Prüfungsforschung haben wir auf unterscheidbare Funktionen von Prüfungen hingewiesen. Wir wollen die dort abgeleiteten Hypothesen noch einmal wiederholen:

- Hypothese 1: Kandidaten in Abschlußprüfungen erreichen im Schnitt bessere Zensuren als Kandidaten in Zwischenprüfungen.
- Hypothese 2: Die Durchfallquote in Zwischenprüfungen ist höher.
- Hypothese 3: Kandidaten in Zwischen- und Abschlußprüfungen unterscheiden sich in konkret angebbaren Strategien und Verhaltensweisen.

Nun können wir einen inherenten Punkt jeder Studienordnung zur Integration der LH-Theorie aufgreifen, nämlich daß Kandidaten in Abschlußprüfungen mit Sicherheit ihre Zwischenprüfungen abgelegt haben, während Kandidaten in Zwischenprüfungen wahrscheinlich weniger Vorerfahrungen mit Prüfungen haben als Abschlußkandidaten. Dies entspricht dem Paradigma der Immunisierung (vgl. S. 63). Es besagt in anderer Formulierung, daß mit der Gelegenheit, kontrollierende Vorerfahrungen zu sammeln, die Wahrscheinlichkeit von Kontrollstrategien wächst. Da Kontrollstrategien das Gegenteil der Strategie der Gelernten hilflosig-

keit darstellen, können wir - indem wir die Wahrscheinlichkeitsaussage in eine Aussage zu Gruppenunterschieden überführen - die nächste Hypothese ableiten:

- Hypothese 5: Kandidaten in Zwischenprüfungen realisieren häufiger die Strategie der Gelernten Hilflosigkeit als Kandidaten in Abschlußprüfungen.

Wir haben somit eine Hypothesentriade vor uns, die wir in Abbildung 5 noch einmal zusammenfassen.

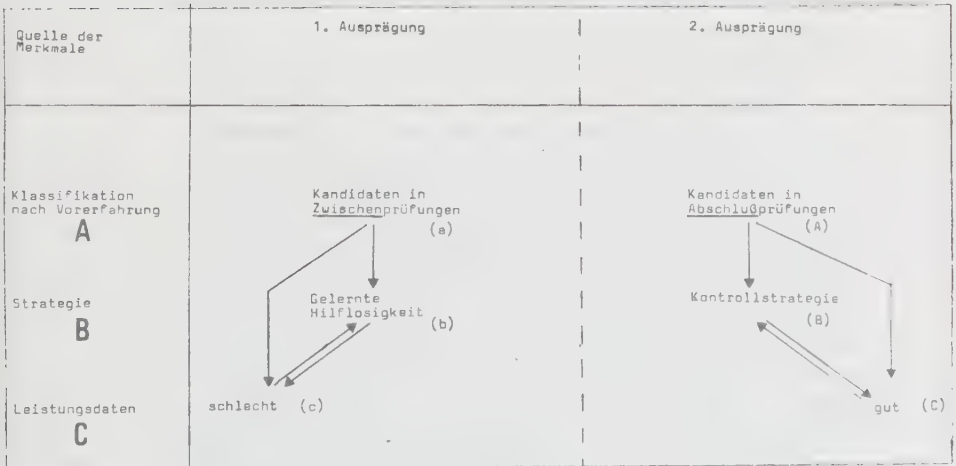


Abbildung 5: Hypothesentriade zu Vorerfahrungen, Strategien und Prüfungsergebnissen

Die Pfeile AC und ac entsprechen dabei den Hypothesen 1 und 2. Der Kontrast AB vs. ab entspricht den Hypothesen 3 und 5. Die Entsprechung von Hypothese 4 findet sich in den Pfeilen CB und cb bzw. BC und bc.

Offen bleibt an dieser Stelle, was wir uns denn nun unter der Strategie der Gelernten Hilflosigkeit vorzustellen haben bzw. was genau eine 'Kontrollstrategie' ist. Ohne dies näher zu spezifizieren, können wir dennoch Beziehungen dieses Konstruktes zu anderen Variablen ableiten.

Bislang undiskutiert blieb die Beziehung der Strategien zum emotionalen Defizit, der Prüfungsbelastung. Im folgenden Kapitel haben wir die drei Betrachtungsebenen (organisch-physiologisch, motorisch-verhaltensmäßig und subjektiv) zusammengefaßt.

2.3.2 Organisch-physiologische, motorisch-verhaltensmäßige und subjektive Betrachtungsebenen

2.3.2.1 Die Erklärung der Depression durch die Theorie der Gelernten Hilflosigkeit

SELIGMAN (1973, 1974, 1975) schlug vor, das 'Learned Helplessness'-Paradigma zur Erklärung von reaktiven Depressionen (SELIGMAN, 1973, 1974) bzw. aller depressiven Erscheinungsformen heranzuziehen (SELIGMAN, 1975).

Da FRESE & SCHÖFTHALER-RÜHL (1976) eine ausgezeichnete deutsche Fassung dieses Ansatzes vorgelegt haben, sei er hier nur kurz dargestellt.¹

2.3.2.1.1 Passivität, kognitives Set und Zeitverlauf

SELIGMAN nannte als vergleichbare Symptome zunächst Passivität und einen negativen kognitiven Set.

Passivität wird in der LH-Forschung in der Regel als verminderte Frequenz einer Kriteriumsreaktion operationalisiert (vgl. z. B. HIROTO, 1974; HIROTO & SELIGMAN, 1975). Passivität als Merkmal depressiven Verhaltens wird von FERSTER (1973) betont (vgl. auch MILLER, 1975).

Der negative kognitive Set zeigt sich im LH-Experiment am besten bei KLEIN et al. (1976); sie wiesen nach, daß Hilflosigkeitstraining und Depressivität zu einem gleichartigen Leistungsdekrement führten. MILLER & SELIGMAN (1973) und MILLER, SELIGMAN & KURLANDER (1975) zeigten, daß Depressive von Versuchskontingenzen weniger profitierten als Nicht-depressive, und interpretieren dies als Unfähigkeit, Kontingenzen zu lernen. Depressive zeigten damit

¹Nach Fertigstellung des Manuskripts und kurz vor Abgabetermin dieser Arbeit stießen wir auf einige - z. T. kritische - weiterführende Artikel im Sonderband des "Journal of Abnormal Psychology" (1978, Volume 87, Number 1) zu diesem Thema.

ein ähnliches Verhalten wie Versuchspersonen mit external-locus-of-control (JAMES & ROTTER, 1958; PHARES, 1957; ROTTER, LIVERANT & CROWEN, 1961). Die Parallelität der Effekte von 'unkontrollierbaren' Vorerfahrungen und einer externalen Kontrollorientierung hat HIROTO (1974) nachgewiesen. HARROW & FERRANTE (1969), WILLIAMS & NICKELS (1969) und ABRAMOWITZ (1969) berichten über eine positive Assoziation zwischen depressiven Symptomen und externaler Orientierung.

BECK (1967) sieht den negativen kognitiven Set als zentrales Symptom der Depression: der Depressive sieht sich, die Umwelt, vor allem die Zukunft negativ, er nimmt vieles gar nicht erst in Angriff, weil er alles für aussichtslos hält. Auch COSTELLO (1972) sieht bei Depressiven einen negativen kognitiven Set, der zu einem "loss of reinforcer effectiveness" führt.

Der nächste gemeinsame Punkt von Gelernten Hilflosigkeit und Depression ist nach SELIGMAN das spontane Abklingen der Effekte und Begleiterscheinungen, sobald eine gewisse Zeit verstrichen ist.

Die spontanen Remissionen von Depressionen sind z. T. Grund gewesen, die sog. 'reaktiven' Depressionen gegen andere Depressionsformen abzugrenzen (vgl. z. B. SCHNEIDER, 1971). SELIGMAN (1973, 1974) hat die LH-Theorie zunächst auch nur auf reaktiv Depressionen übertragen und sie gegen 'endogene' Depressionen mit hereditären biochemischen Ursachen klar abgegrenzt. Später (1975) aber gab er diese explizite Unterscheidung auf und gab zu erkennen (vgl. SELIGMAN, 1975, S. 94), daß er nunmehr einem Kontinuummodell zuneigt, wonach zwischen dem reaktiven und dem endogenen Pol ein Kontinuum der 'An-

fälligkeit' für den Glauben an die Unabhängigkeit von Reaktion und Ergebnis besteht. Reaktiv-Depressive unterscheiden sich demnach von Endogen-Depressiven durch ihre geringere 'Anfälligkeit'.

Ein Schwinden des LH-Effektes im Laufe der Zeit konnte im Humanversuch bislang nicht beobachtet werden; THORNTON & POWELL (1974) fanden nach vierundzwanzig Stunden einen unveränderten Effekt. OVERMIER & SELIGMAN (1967) fanden den einzigen bislang beobachteten zeitlichen Schwund bei Menschen nach achtundvierzig Stunden. Da diese beiden Untersuchungen sich einsam gegenüberstehen, muß die Existenz eines zeitlichen Verlaufs offen bleiben. Trauer nach einem Todesfall als klassisches Beispiel für reaktive Depression kann bekanntlich Wochen anhalten, bevor sie sich mildert. Untersuchungen, die einen längerfristigen Verlauf des LH-Effektes erfassen, stehen noch aus.

Wir können festhalten, daß Passivität und negatives kognitives Set sowohl im LH-Experiment als auch bei Depressiven beobachtet werden konnte. Die Behauptung eines analogen zeitlichen Verlaufs entbehrt jedoch bislang jeder empirischen Grundlage.

Passivität und negativen kognitiven Set haben wir auch schon im Zusammenhang mit Prüfungsangst beschrieben. Prüfungsängstliche verhalten sich vorsichtig und wollen eher wenig sagen, erleben die Prüfung als über sie hereinbrechend und sich als ohnmächtig ausgeliefert.

2.3.2.1.2 Mangelnde Aggressivität

Als nächstes gemeinsames Merkmal zwischen Depression und LH werden fehlende Aggressionen genannt.

Imponierend an Depressionen ist auch für den Laien das fehlende assertive Verhalten. Im sozialen Bereich ist Aggressivität die spektakulärste (und evtl. wirksamste) Wirkreaktion.

ZUNG (1971) nimmt fehlende (offene) Aggression als Kriterium zur Unterscheidung zwischen Angst und Depression.

KURLANDER, MILLER & SELIGMAN (zit. n. SELIGMAN, 1975) wiesen nach, daß in einem Gefangenen-Dilemma-Spiel sich Depressive und 'Hilflose' in ähnlicher Weise häufiger zurückzogen und deshalb verloren.

Andererseits verweisen SELIGMAN (1975) und FRESE & SCHÖFTHALER-RÜHL (1976) auf eine eher passive Hostilität von Depressiven. Unklar bleibt bei den Ausführungen der Autoren, was unter 'Hostilität' genau zu verstehen ist. Manche Depressive fallen durch Verbitterung, Ressentiments und Gereiztheit auf (vgl. SCHULTE & TÖLLE, 1973, S. 66); dies könnte mit 'Hostilität' gemeint sein.

Wenn wir Aggression als eine von mehreren Möglichkeiten interpretieren, aktiv soziale Verstärker zu kontrollieren, dann wäre fehlende Aggression als Teil eines allgemeinen Defizits aktiven Verhaltens zu interpretieren (vgl. EASTMAN, 1976, S. 278 f.).

Dies stünde in Übereinstimmung mit dem LH-theoretischen motivationalen Defizit. Hostilität wäre dann der Ausdruck allgemein passiven offenen Verhaltens.

Dies könnte im LH-Experiment ebenfalls der Fall sein. ROTH & KUBAL (1975) jedenfalls berichten über eine Zunahme des skalierten Ärgers bei ihren 'hilflosen' Gruppen. KLEIN, FENCIL-MORSE & SELIGMAN (1976) erwähnen ohne nähere Angaben ähnliche Daten. Zur gleichen Zeit berichten ROTH & KUBAL (1975) ebenso wie TENNEN & ELLER (1977) über verstärkten Ärger bei 'reaktanten' Gruppen. Eine Studie, die sich explizit mit Aggressionen unter unkontrollierbaren Bedingungen beschäftigt, steht jedoch noch aus.

Auch in Studien zur Prüfungsangst finden wir zuweilen Hinweise auf aggressive Komponenten. HEINERTH (1972a) interpretiert einen der von ihm gefundenen Faktoren mit 'Aggressivität/Frustrabilität'. HECKHAUSEN (1977) fand 'Verärgerung' im Zusammenhang mit einem ungünstigen Prüfungsverlauf. Die weit überwiegende Mehrheit der Autoren, die zum Thema Prüfungsangst veröffentlicht haben, erwähnen jedoch keine Verhaltensweisen, die man als Aggressivität oder Hostilität bezeichnen könnte.

LH-Effekt, depressives Verhalten und Prüfungsangst scheinen sich demnach in fehlendem aggressiven Verhalten zu ähneln, wenngleich eine klare Aussage wegen der geringen Datendecke z. Z. noch nicht möglich ist. Evtl. reduziert sich das Problem auf die bereits konstatierte ähnliche Passivität.

2.3.2.1.3 Sexuelle, soziale und Ernährungs- defizite, Störungen des Noradrenalin- haushalts und Magengeschwüre

Als weitere Punkte der Ähnlichkeit zwischen LH-Effekt und depressivem Verhalten nennt SELIGMAN:

- sexuelle Defizite (Impotenz, verringertes sexuelles Interesse),
- soziale Defizite (Interessenverlust),
- Gewichtsverluste,
- Störungen des Noradrenalinhaushalts und
- Magengeschwüre.

Die Ähnlichkeit wird mit entsprechenden Befunden aus LH-Tierexperimenten begründet. Für keinen der obigen Punkte finden sich Belege aus Humanexperimenten, auch nicht in Form von Analogiestudien (man könnte z. B. versuchen, sexuelle Defizite der UK-Gruppe durch verringertes physiologische Reaktionen auf sexuelles Reizmaterial nachzuweisen o. ä.)

Die fünf von SELIGMAN angegebenen Punkte sind demnach weniger Parallelen, sondern zeigen Forschungsgebiete, in denen Parallelen erst noch nachgewiesen werden müssen.

Wenn wir die fünf Punkte auf das Gebiet der Prüfungsangst übertragen, finden wir ebenfalls wenig empirische Daten. Zwar fand HEINERTH (1972a) einen Faktor, den er mit 'Essstörungen' interpretierte, was ein Ernährungsdefizit andeuten könnte, und

SCHEER & ZENZ (1973) fanden einen Faktor 'Selbstsicherheit', dessen negativen Pol wir als soziales Defizit auffassen könnten, jedoch gibt es darüber hinaus keine empirischen Parallelen. Auch hier finden wir daher überwiegend Forschungsanregungen zur Feststellung der Symptomatik von Prüfungsbelasteten.

2.3.2.1.4 Affekte

Gefühle der Hilflosigkeit und Inkompetenz sind zentrale Affekte der affektiven Störung Depression (vgl. SCHULTE & TÖLLE, 1973). Für deren Wirksamkeit oder Auftreten im LH-Experiment finden wir allenfalls Indizien (ROTH & KUBAL, 1975; KLEIN, FENCIL-MORSE & SELIGMAN, 1976).

Das Gefühl der Hilflosigkeit und die 'Gelernte Hilflosigkeit' sind von der physiologischen, motorischen und kognitiven Maebene her nicht zu unterscheiden. Der Unterschied liegt in der theoretischen Konzeption, wonach Gelernte Hilflosigkeit ein Trait-System ist (HIROTO & SELIGMAN, 1975, S. 326). Daraus kann gefolgert werden, da das Gefühl der Hilflosigkeit eine aktuelle, eher kurzfristige state-Emotion ist (vgl. SPIELBERGER, 1966, zur Trait-State-Unterscheidung).

Als weiteren gemeinsamen Affekt von LH und Depression nennt SELIGMAN die Angst. Im LH-Humanexperiment wurde Angst bislang nicht nachgewiesen.

Die Ähnlichkeit der Affekte zwischen Gelernter Hilflosigkeit und Depression konnte bislang demnach nicht überzeugend gezeigt werden.

Abgesehen von diesem Punkt läßt sich dennoch zusammenfassend sagen, daß LH und Depression deutlich ähnliche Reaktionskomponenten haben. Zwar sind die Belege für einige der von SELIGMAN genannten Punkte nicht unbedingt überzeugend - hier muß noch viel Forschungsarbeit getan werden -, wir haben aber keine Veröffentlichung gefunden, deren Ergebnisse eindeutig gegen die LH-theoretischen Erklärung der Depression sprechen.

Damit verwenden wir jedoch eine Theorie zur Erklärung von 'depressiven Reaktionen' und übertragen sie auf den Bereich der Prüfungs-'Angst'.

Zwar haben wir bereits auf das typische Angst-/Depressionsgemisch Prüfungsbelasteter hingewiesen, es stellt sich dabei jedoch die grundsätzlichere Frage nach dem Verhältnis von Angst und Depression.

In dem folgenden Exkurs werden wir die von uns gefundenen empirischen Ergebnisse zum Verhältnis von Angst und Depression darstellen.

2.3.2.2 Exkurs: Zum allgemeinen Verhältnis von Angst und Depression

Intuitiv haben Angst und Depression wenig gemeinsam, beziehen sich auf scheinbar völlig verschiedene Zustände, Emotionen. Im Gegensatz dazu stehen die Ergebnisse der bisherigen Forschung, in denen eine Unterscheidung von 'Angst' und Depression' nicht deutlich gelang.

SCHULTE & TÖLLE (1973) unterscheiden beispielsweise zwischen Angstneurose und neurotischer Depression. Wie bei allen neurotischen Störungen tritt Angst auch bei der depressiven Reaktion auf; wird das Krankheitsbild jedoch ausschließlich von Angst begleitet, spricht man von Angstneurose. Bei medikamentöser Therapie sind bei Angstneurotikern z. B. Tranquilizer, bei Depressiven z. B. Thymoleptika indiziert, zur vorübergehenden pharmakologischen Entspannung evtl. aber auch wieder Tranquilizer.

MARKS (1969) berichtet, daß Agoraphobiker - MARKS benutzt den Ausdruck 'panphobic syndrom' - häufig von Begleiterscheinungen der Erschöpfung und Depression berichten. MARKS erwähnt das Konzept der 'burnt out habits', mit dem die fehlende oder geringe Therapierbarkeit während akuter Phasen erklärt werden könnte.

MILLER, SELIGMAN & KURLANDER (1975) berichten von Korrelationen von Fragebogenscores bei Studenten in Höhe von $r = .41$ ($p < .01$). MENDELS et al. (1972) nennen Korrelationen zwischen Angst- und Depressionsskalen zwischen $r = .50$ und $r = .70$ bei Patienten. SCHLESS et al. (1974) fanden eine Korrelation zwischen dem BECK-Depressions-Inventory und einer Angstsymptomcheckliste von $r = .66$ bei Patienten.

Die meisten der Faktorenanalysen, in denen die Daten depressiver Patienten verrechnet wurden, ergaben einen Faktor, der mit 'Angst' bezeichnet werden kann, sofern nicht Angst und Depression zusammenfielen (vgl. zusammenfassend MENDELS et al., 1972).

ZUCKERMAN, PERSKY, ECKMAN & HOPKINS (1967) untersuchten nach der Logik von CAMPBELL & FISKE (1959) Depressive, Angstneurotiker und Normale mit Fremdbeurteilungs-, verschiedenen Selbstbeurteilungs- und projektiven Verfahren. Sie berichten von guter konvergenter Validität der Meßverfahren für Angst und Depression, aber schlechter diskriminanter. Eine Inspektion ihrer Daten ergibt, daß die konvergenten Validitätsmaße in etwa der gleichen Höhe liegen wie die diskriminanten ($r = .40$ bis $r = .70$).

MILLER, SELIGMAN & KURLANDER (1975) teilten freiwillige studentische Versuchspersonen in ihrem Ex-

periment zum einen nach dem BECK-Depressions-Inventory (BECK, 1967), zum anderen nach der IPAT-Anxiety-Scale (CATTELL & SCHEIER, 1961) auf. Ihr ursprünglicher Plan, vier Gruppen (größer bzw. kleiner als der Mittelwert der beiden Tests) zu bilden, konnte nicht realisiert werden. Versuchspersonen, die als unterdurchschnittlich ängstlich und überdurchschnittlich depressiv zu klassifizieren waren, konnten nicht in ausreichender Zahl gefunden werden.

ZUNG (1971) untersuchte 225 Neuaufnahmen einer psychiatrischen Klinik mit je einer Selbstbeurteilungsskala für Depression und Angst (ZUNG, 1965, 1971), die beide klinisch validiert worden waren, und einer Fremdbeurteilungsskala, die Angst und Depression trennen soll (DORFMAN & LINDNER, 1958, zit. n. ZUNG, 1971).

Im Fremdbeurteilungsfragebogen werden fünfzehn Bereiche erfaßt und dabei jeweils ängstliche und depressive Verhaltensweisen einander gegenübergestellt. Beispielsweise gilt als ängstliche Stimmung "anxious, worried, tense, nervous", als depressive "sad, blue, down-hearted, depressed"; ängstliche Sprache und Sprachtönung werden mit "rapid, tense, excited, high-pitched", depressive mit "slow, inanimated, tearful, mask-like" beschrieben; Gestik und Auftreten Ängstlicher werden die Merkmale "tense, tendency to shift position frequently" zugeordnet, bei Depressiven heißt es "droopy head and shoulders, limp, draggy"; für Ängstliche steht das Begriffspaar "difficulty in falling asleep, nightmares", für Depressive "repeated awakening, early morning awakening"; der Antrieb und die Energie Ängstlicher werden mit "accelerated drive, spurts of energy - then exhaustion", der Depressiver mit "lack of drive, energy drained, fatigued" gekennzeichnet. Demnach scheint die DORFMAN-LINDNER-Skala - der einzige Test, in dem Angst und Depression als Gegensätze konzipiert sind - Ängstliche und Depressive nach ihrer Bewegungsfrequenz und/oder -latenz zu unterscheiden.

ZUNG berichtet keinerlei Korrelationen und interpretiert das Ergebnis seiner Gruppenvergleiche wie folgt:

"Aus den Resultaten dieser Studie ergibt sich ein dritter alternativer Standpunkt zu Angst und Depression als klinischen Einheiten; während Angst mit

Depression in einem signifikanten Ausmaß assoziiert ist (wie sich aus den Ergebnissen der DORFMAN-LINDNER-Skala ergibt) und vielleicht als Angst-Depressions-Symptom Komplex betrachtet werden sollte, kann Depression als einzelne und abgegrenzte Einheit auftreten (wie sich aus den Ergebnissen der getrennten Angst- und Depressions-Ratingskala ergibt)." (ZUNG, 1971, S. 383, übersetzt vom Verfasser)

Zu ähnlichen Schlußfolgerungen gelangen MENDELS et al. (1972) und ZUBIN & FLEISS (1970) nach Analyse ihrer Patientendaten.

KERR, SCHAPIRA, ROTH & GARSIDE (1970) verglichen Gruppen von Agoraphobikern, Patienten mit spezifischen Phobien, Reaktiv-Depressiven und Endogen-Depressiven und verwendeten dazu Verlaufswerte des MPI (MAUDSLEY Personality Inventory, EYSENCK, 1959). Übereinstimmend mit SNAITH (1968) und MARKS & GELDER (1965) beobachteten die Autoren bei Agoraphobikern ein "depressive colouring", was wieder einmal ein gemeinsames Auftreten von ängstlichen und depressiven Reaktionen andeutet.

SCHAPIRA, KERR & ROTH (1970) untersuchten Patienten, die zum einen als ängstliche Neurotiker, zum anderen als depressive Neurotiker diagnostiziert worden waren, hinsichtlich der Häufigkeit von Phobien vor, während und nach der akuten Erkrankung. Ihre Ergebnisse lassen sie vermuten, daß Patienten, die später an einer Angstsymptomatik leiden werden, vor und während der Erkrankung eine höhere Anzahl von Phobien haben als Patienten, die später ein depressives Zustandsbild entwickeln. 'Ängstliche' sind zudem introvertierter.

Zusammenfassend können wir sagen, daß Angst und Depression keineswegs distinkte Entitäten sind, sondern daß sie sehr häufig gemeinsam auftreten. Auch scheinen sie nicht unabhängig voneinander zu sein, wenngleich wir kaum falsifizierbare Aussagen über ihren Zusammenhang finden.

Der gegenwärtige Stand der Forschung über das Verhältnis von Angst und Depression ist - wie wir versuchten aufzuzeigen - noch rudimentär.

Für diese Arbeit begnügen wir uns mit der A-priori-Hypothese, daß Angst und Depression noch korrelieren

und sich zur Prüfungsbelastung zusammenfinden (Hypothese 6).

Forschungsfragestellungen, die sich aus der Unterscheidung von Angst und Depression ergeben, wollen wir in dieser Arbeit ausklammern.

2.3.2.3 Ableitung der Hypothesen zum Verhältnis von Vorerfahrungen, Prüfungsstrategien und Prüfungsbelastungen

Das emotionale Defizit im LH-Humanexperiment ist bislang kaum erforscht. Insofern, als die Theorie der Gelernten Hilflosigkeit auch depressives Verhalten erklärt, können wir mit Bezug auf den vorhergehenden Exkurs und mit Verweis auf die empirischen Ergebnisse zur Prüfungsbelastung die folgenden Hypothesen ableiten. Wir legen dieselbe Gedankenführung zugrunde wie bei der Ableitung der Hypothesen zu den Leistungsdaten.

Zunächst Hypothese 6, die die Messung der Prüfungsbelastung betrifft: die Indikatoren von Angst und Depression korrelieren hoch positiv und bilden gemeinsam die Symptomatik der Prüfungsbelastung.

Hypothese 7: Kandidaten, die die Strategie der Gelernten Hilflosigkeit realisieren, sind stärkeren Prüfungsbelastungen unterworfen als Kandidaten, die Kontrollstrategien einsetzen.

Hypothese 8: Kandidaten in Zwischenprüfungen leiden unter stärkeren Prüfungsbelastungen als Kandidaten in Abschlußprüfungen.

Wir wollen die Hypothesen 6, 7 und 8 nun wieder zu einer Hypothesentriade zusammenfassen.

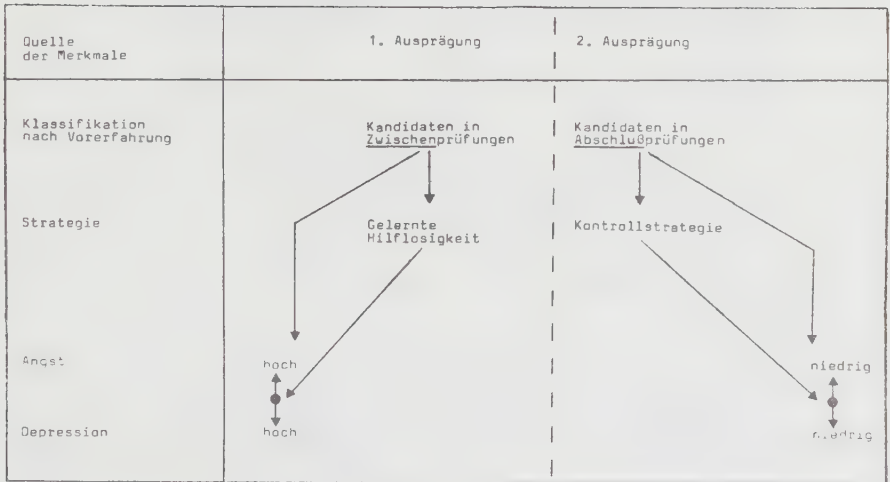


Abbildung 6: Hypothesentriade zu Vorerfahrungen, Prüfungsstrategien und Prüfungsbelastungen

An dieser Stelle wollen wir darauf hinweisen, daß die Hypothese 8 im Widerspruch zur naiven Auffassung steht, Prüfungsangst sei die Folge des Risikos durchzufallen. Der Klarheit wegen wollen wir diese Auffassung zunächst umformulieren: Je geringer die Wahrscheinlichkeit ist, ein 'Ungut' oder 'Mangelhaft' zu erhalten, desto geringer ist die Prüfungsangst. Erst die Zusatzannahme der unterschiedlichen Selektionsbedingungen kann zu der Hypothese 8 führen, keinesfalls jedoch zu den Hypothesentriaden.

Noch ein weiterer Unterschied besteht zwischen der naiven Auffassung und unserer Konzeption: Wir gehen davon aus, daß die Kandidaten Strategien einsetzen, um Operationen herauszufinden, mit denen sie ein möglichst gutes Ergebnis erreichen können; selbst wenn ein Kandidat sicher sein kann, die Prüfung zu bestehen, wird er doch von Angst geschüttelt, wenn er die Strategie der Gelernten Hilflosigkeit einsetzt. Diese relative Unabhängigkeit der Prüfungsbelastung von dem Durchfallkriterium kann die o. a. naive Auffassung nicht erklären.

2.3.3 Integration der Betrachtungsebenen: Beziehungen zwischen Vorerfahrungen, Prüfungsstrategien, Leistungsdaten und Prüfungsbelastungen

Wir wollen nun die beiden Hypothesentriaden vereinigen.

Zunächst ist die Frage nach dem allgemeinen Verhältnis von Prüfungsbelastung und Leistungsdaten zu beantworten. Hierbei ist zu berücksichtigen, daß wir es nicht mit einfachen kognitiven Leistungen wie Rechenaufgaben, Buchstabenfolgen usw. zu tun haben, sondern mit sehr komplexen. Während für einfache Aufgaben im allgemeinen eine umgekehrt-u-förmige Beziehung zwischen Leistung und Emotionen (= Aktivierung) angenommen werden kann (vgl. SCHÖNPFLUG, 1969; McLEAN, 1967; CORTEEN, 1969; TEICHNER, 1968, u. v. a.), gibt es Autoren, die zwischen komplexen kognitiven Leistungen und Emotionen keine (additive) Beziehung sehen (z. B. AUSUBEL, 1974, S. 452), insbesondere, wenn Noten als Leistungsindex benutzt werden (vgl. SPIELBERGER, GORSUCH & LUSHENE, 1970).

Wir nehmen an, daß zwischen den Leistungen in mündlichen akademischen Prüfungen und Prüfungsbelastungen keine einfache Beziehung besteht, sondern eine, die einer Partialkorrelation entspricht (bei additiven Beziehungen), oder eine andere mit komplexen Beziehungen höherer Ordnung (die wir noch nicht kennen). Für die Beziehungen zwischen Belastungs- und Leistungsindikatoren formulieren wir deshalb keine spezifizierten A-priori-Hypothesen.

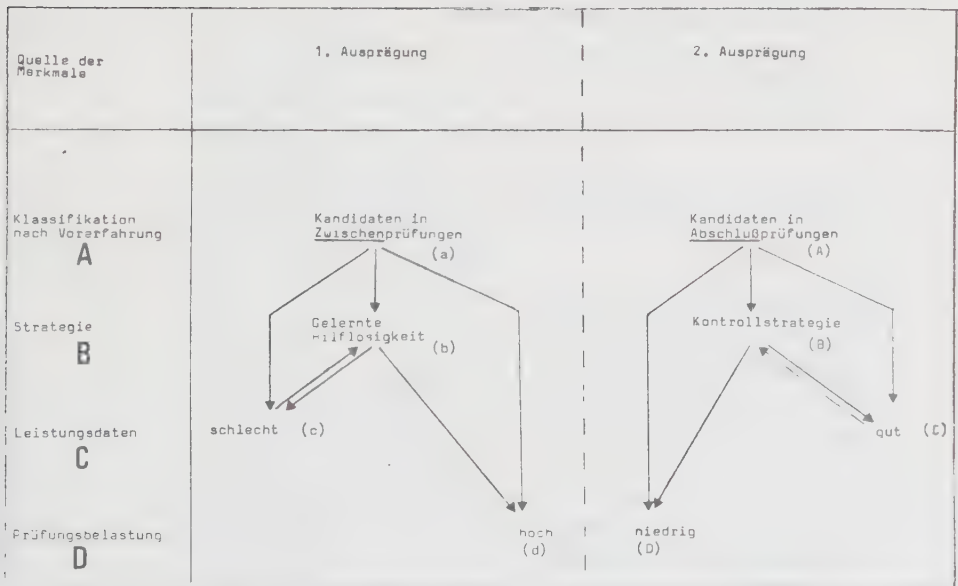


Abbildung 7: Beziehungen zwischen Vorerfahrungen, Prüfungsstrategien, Leistungsdaten und Prüfungsbelastungen

In Abbildung 7 haben wir versucht, die Beziehungen graphisch darzustellen. Die Pfeile stellen statistische Prüfoperationen und logische Abhängigkeiten dar. Im folgenden führen wir die bislang abgeleiteten Hypothesen noch einmal im Wortlaut und mit Bezug auf Abbildung 7 formalisiert auf. Dabei stellt '+' das formale Zeichen für eine Zusammenhangshypothese und 'vs.' das Zeichen für eine Kontrastaussage dar.

Hypothesen:

Formalisierung:

Hypothese 1: Kandidaten in Abschlußprüfungen erhalten im Schnitt bessere Zensuren als Kandidaten in Zwischenprüfungen.

$A + C \text{ vs. } a + c$

Hypothese 2: Die Durchfallquote in Zwischenprüfungen ist höher.

$a + c \text{ vs. } A + C$

Hypothese 3: Kandidaten in Zwischen- und in Abschlußprüfungen unterscheiden sich in konkret angebbaren Strategien und Verhaltensweisen.

$A + B \text{ vs. } a + b$

Hypothese 4: Erfolgreiche Kandidaten realisieren aktive und kontrollierende Strategien, während nicht-erfolgreiche passiv und hilflos agieren.

$C + B \text{ vs. } c + b$

Hypothese 5: Kandidaten in Zwischenprüfungen realisieren häufiger die Strategie der Gelernten Hilflosigkeit als Kandidaten in Abschlußprüfungen.

$a + b$ vs. $A + b$

Hypothese 6: Die Indikatoren von Angst und Depression korrelieren hoch positiv und bilden gemeinsam die Symptomatik der Prüfungsbelastung.

D vs. d

Hypothese 7: Kandidaten, die die Strategie der Gelernten Hilflosigkeit realisieren, sind stärkeren Prüfungsbelastungen unterworfen als Kandidaten, die Kontrollstrategien einsetzen.

$b + d$ vs. $B + D$

Hypothese 8: Kandidaten in Zwischenprüfungen leiden unter stärkeren Prüfungsbelastungen als Kandidaten in Abschlußprüfungen.

$a + d$ vs. $A + D$

Offen blieb bis hierhin, was genau wir uns unter Prüfungsstrategien vorzustellen haben. In der Einleitung schrieben wir, daß wir hierfür eine Mischung aus Hypothesengenerierung und Hypothesenprüfung als Analogie zur Konstruktvalidierung vorgesehen haben. Im folgenden Kapitel gehen wir auf die Möglichkeiten der Messung von Prüfungsstrategien ein. Indem wir diesem Kapitel vorgreifen und die formalisierte Hypothese (B vs. b) formulieren, gelangen wir zu folgender Reihenfolge der Hypothesenprüfungen und -generierungen:

1. $A + C$ vs. $a + c$ (Hypothesen 1 + 2)
2. D vs. d (Hypothese 6)
3. $A + D$ vs. $a + d$ (Hypothese 8)
4. B vs. b (im folgenden Kapitel)
- — — — — — — — —
5. $C + B$ vs. $c + b$ (Hypothese 4), zugleich
 $B + c$ vs. $b + c$
6. $A + B$ vs. $a + b$ (Hypothese 3 u. 5)
7. $B + D$ vs. $b + d$ (Hypothese 7), zugleich
 $D + B$ vs. $d + b$
- — — — — — — — —
8. $A + B + C + D$ vs. $a + b + c + d$ (?)

2.3.4 Prüfungsstrategien: ihre Messung und Valididierung

Wir wollen das bisher Geschriebene zusammenfassen:

Wir haben geschrieben, daß Prüfungen für den Kandidaten ein Problem darstellen, bei dem Ausgangssituation und Zielzustand definiert sind und bei dem der Kandidat die Aufgabe hat, Operationen zu finden, um Ausgangssituation und Zielzustand zu verknüpfen. Wir haben im sozialwissenschaftlichen Teil die Funktionen der Prüfungen bestimmt und von daher Bestimmungsmerkmale für das Problem Prüfung abgeleitet: Ausgangspunkt ist ein abgeleitetes Studium und eine inferiore Position, Zielzustand ist eine möglichst gute Zensur und eine symmetrische Beziehung zum Prüfer.

Um die beiden Bestimmungsstücke miteinander verknüpfen zu können, muß der Prüfling in das Kommunikationssystem Prüfung aktiv eingreifen und aus einer theoretisch unendlichen Menge von Operationen die geeigneten auswählen. Um den Suchprozeß nicht dem Zufall zu überlassen, benutzt der Prüfling Strategien. Strategien umfassen Erwartungen und Taktiken und bestimmen das Verhalten in der Prüfung.

Als eine Möglichkeit von Strategien hatten wir die Gelernte Hilflosigkeit kennengelernt, deren Effekte als kognitives, motivationales und emotionales Defizit beschrieben werden. Wir haben diese Theorie dargestellt, sie auf die Situation der

Prüfung übertragen und daraus Hypothesen abgeleitet.

Ein LH-Effekt als Gruppenunterschied kann mit vier Methoden nachgewiesen werden:

1. Klassifikation der Teilnehmer in Gruppen mit unterschiedlicher Vorerfahrung,
2. direkte Messung der Gelernten Hilflosigkeit,
3. Realisation des LH-Effektes durch Trainingsphasen im Experiment,
4. unterschiedliche Instruktion.

Da unser Hintergrundinteresse in der Prophylaxe echter Prüfungsbelastungen besteht, entfallen die dritte und die vierte Möglichkeit. Die Klassifikation der Teilnehmer ist dann möglich, wenn wir die Variable 'Prüfungserfahrung' berücksichtigen. Zur direkten Messung der Gelernten Hilflosigkeit in der Prüfung haben wir noch nicht Stellung bezogen. Dies soll im folgenden geschehen.

Wir haben Gelernte Hilflosigkeit als Strategie bezeichnet und als Gegensatz den Begriff der 'Kontrollstrategie' konzipiert.

Die bisherigen Untersuchungen zur Funktion von Strategien beim Begriffsbilden und beim Problemlösen (vgl. KLIX, 1973) sind u. E. auf die Situation des Prüfungskandidaten nicht anzuwenden. Die Mehrheit der Veröffentlichungen verwendet Probleme, in denen es eine optimale Sequenz von Operationen gibt, mit hin jede Abweichung vom Zielzustand wegführt. Zudem handelt es sich meist um Probleme, deren Struktur mathematisch vollständig beschrieben werden kann.

Wir sind nicht der Auffassung, daß es eine optimale Sequenz von Operationen für mündliche Prüfungen gibt. Unterschiede zwischen einzelnen Fächern (und evtl. Prüfern) sollten hier durchaus zum Tragen kommen. Wir sind aber der Auffassung, daß alle optimalen Strategien für mündliche Prüfungen einen gemeinsamen Teil haben - die Kontrollstrategie - und daß die Negation der Kontrollstrategie die Strategie der Gelernten Hilflosigkeit ist.

Zugleich wissen wir, daß es einige Operationen geben muß, die sowohl in Beziehung zu schlechten Zensuren wie in Beziehung zu Prüfungsbelastungen stehen (vgl. S. 40, z. B.: vorsichtig und selbstkritisch auftreten, wenig sagen, ggf. nicht widersprechen usw.). Die verfügbaren Anhaltspunkte sind jedoch nicht detailliert genug, um letztlich die Hoffnung auf Prüfungsbelastungsprophylaxe zu erfüllen. Was wir suchen, ist ein Inventar konkreter Strategien.

Methodisch gesehen bedeutet dies, daß wir sowohl ein Meßinstrument suchen als auch seine Validität nachweisen wollen. Wir haben uns deshalb zu einer vom Üblichen abweichenden sequentiellen Vorgehensweise entschlossen. In einem Suchprozeß werden wir zunächst nach geeigneten Meßinstrumenten fahnden, um diese dann auf Validität zu überprüfen. Der wichtigste Gesichtspunkt ist dabei der der Konstruktvalidität, d. h. der Nachweis, daß die von uns gefundenen 'Strategien' in genau dem Zusammenhang mit den Hypothesen stehen, den wir formuliert haben (vgl. DOHRENDWEND & DOHRENDWEND, 1965).

Dies ist der eine Weg, den wir gehen können: Wir stellen eine Sammlung potentieller Strategien zusammen und schauen, welche dieser Strategien in dem oben beschriebenen theoretischen Zusammenhang stehen und ob sie inhaltlich dem bislang Geschriebenen entsprechen.

Inhaltlich-methodisch stellt sich hierbei die Frage, wie wir uns Strategien vorzustellen haben. Sie können Dimensionen im Sinne des faktorenanalytischen Modells darstellen, aber auch adimensionalen Strukturen entsprechen (... wenn der Prüfer ..., dann ... ich ..., wenn dann ...). Dabei stellt sich aber nicht nur das Problem der Gruppierung von Items (wir gehen darauf im methodischen Teil ein), sondern auch das der 'Strukturiertheit' von Strategien. Wir meinen damit, daß Gelernte Hilflosigkeit u. U. kein Kontrast zur Kontrollstrategie im Sinne einer Umkehrung der Vorzeichen ist, sondern sich evtl. darin ausdrückt, daß 'Hilflose' im Gegensatz zu 'Kontrollierenden' Zufallswerte skalieren. Dies würde bedeuten, daß das übliche Verfahren, eine Zusammenstellung potentieller Strategien zu analysieren - die Faktorenanalyse - für uns versperrt wird.

Hier bietet sich die Möglichkeit, die bei der Integration des Konfliktkonzepts herausgearbeitete Implikation aufzugreifen: 'Hilflose' zeichnen sich durch einen Mangel an diskriminativen Reizen aus.

Eine zweite methodische Möglichkeit zur Messung von Strategien liegt in der Abstraktion von der Prüfungssituation.

Wenn wir die Prüfung aus der Sicht des Kandidaten als eine Situation betrachten, in der in erster Linie ein kommunikatives Problem zu lösen ist, dann können wir aus der Literatur Ergebnisse zu (allgemeinen) Problemlösestrategien heranziehen.

ERTEL (1966) und ERTEL & SCHINDLER (1969a, b) legten ein Funktionsmodell des Testverhaltens und zu dessen Messung das Testverhaltensdifferential vor. Das Testverhaltensdifferential wurde bislang im Zusammenhang mit Intelligenztests (ERTEL & SCHINDLER, 1969a, b) und bei der experimentellen Analyse der Feldabhängigkeit (SCHULTE, 1972) verwendet.

Das Testverhaltensdifferential mißt das Testverhalten in drei Dimensionen: mäßige/erhöhte Anspannung, Variabilität/Stetigkeit und Unifizierung/Diversifizierung. Nähere Einzelheiten sind im methodischen Teil aufgeführt.

Auf die funktionalen Analogien zwischen der Situation des Prüfungskandidaten und der Situation der Versuchsperson im Problemlöseexperiment haben wir bereits hingewiesen. SCHULTE (1972) verwendete zwei verschiedene Instruktionen, bei denen es um die Annäherung an einen 'objektiven' Parameter (der senkrechten Stellung eines Stabes) bzw. um den 'subjektiven' ersten Eindruck ging. Die Instruktionen ähnelten funktional der 'chance'- und der 'skill'-Instruktion der Locus-of-Control-Forschung.

Die 'objektive' Instruktion lautete u. a.:

"Wir machen zunächst einen Versuch, bei dem ich feststellen will, wie gut Sie unter verschiedenen Bedingungen die Senkrechte einstellen können. ... so sagen Sie mir bitte, ob der Stab senkrecht steht, d. h. die Schwerkraftrichtung innehat, oder ob das obere Ende des Stabes nach rechts oder nach links geneigt ist. Sie sollen also die Stellung des Stabes angeben, indem Sie die Lage des oberen Endes des Stabes beurteilen... Sobald er objektiv senkrecht ist, sagen Sie bitte "stopp"."
(SCHULTE, 1972, S. 176f.)

Die 'phänomenale' Instruktion lautete u. a.:

"Ich möchte Sie bitten, keine Überlegungen anzustellen. Bitte entspannen Sie sich und beurteilen Sie die Lage des Stabes wirklich nur nach Ihrem ersten Eindruck. Mit geht es lediglich darum festzustellen, wann der Stab für Sie senkrecht aussieht, unabhängig davon, ob das stimmt oder falsch ist. Es gibt also bei diesem Versuch kein richtiges oder falsches Urteil." (SCHULTE, 1972, S. 177)

Wir haben den Eindruck, daß die 'objektive' Instruktion einer 'skill'-Instruktion, die 'phänomenale' einer 'chance'-Instruktion entsprechen könnte, zumindest aber, daß der Kontrast 'phänomenal' zu 'objektiv' dem Kontrast 'chance' zu 'skill' entspricht.

SCHULTE konnte nachweisen, daß die Reaktionen im Testverhaltensdifferential mit den unterschiedlichen Instruktionen variieren und daß ein Zusammenhang der Veränderung der Strategie mit dem Instruktionswechsel besteht.

Wir gehen davon aus, daß das Testverhaltensdifferential auch in der Lage ist, unterschiedliches Vorgehen in Prüfungen abzubilden.

Wenn unsere Vermutung richtig ist, daß SCHULTES 'objektive' Instruktion einer internalen Kontrollorientierung funktional äquivalent ist, dann sollte die Kontrollorientierung für die Prüfung ähnliche Merkmalsausprägungen erreichen: die 'objektive' skill-Instruktion führte zu größerer Anspannung, größerer Stetigkeit und stärkerer Diversifizierung. Wenn die Auseinandersetzung des Prüfungskandidaten der Auseinandersetzung der Versuchsperson mit den Versuchsbedingungen ähnelt, dann sollte eine internalen Kontrollorientierung einer größeren Anspannung, einer größeren Stetigkeit und einer geringeren Unifizierung entsprechen.

Der Unterschied zwischen dem Experiment und der Prüfung liegt darin, daß im Experiment die Kontrollorientierung induziert wurde, während wir für die Prüfung die Kontrollorientierung als Organismusvariable Kontrollstrategie konzipieren. Gemeinsam ist beiden Situationen, daß die von der Versuchsperson/dem Kandidaten in die Situation mit eingebrachte intendierte Kontrollstrategie sich realisieren läßt oder auch nicht, daß zwischen intendierter Kontrolle und realisierter Kontrolle eine Diskrepanz auftreten kann.

Wir wollen beide Möglichkeiten der (mehr oder weniger) direkten Messung von Prüfungsstrategien aufgreifen: die Messung von Strategien für konkrete Prüfungen und - mehr abstrakt - für allgemeine Probleme.

Für beide Datensätze gilt Hypothese 9:

Es gibt mindestens zwei multivariate Profile, wovon das eine einer Kontrollstrategie und das andere der Strategie der Gelernten Hilflosigkeit entspricht und wovon das eine die Umkehrung des anderen ist.

3 METHODISCHER TEIL

3.1 Methodische Rahmenbedingungen:

Die Prüfung als natürliches Experiment

Neben Erwägungen zur statistischen Methodik, der Auswahl und Operationalisierung der relevanten Variablen und der Darstellung der Datenerhebungspraxis werden im methodischen Teil üblicherweise das Design und die daraus ableitbare Stichprobe umfassend erörtert. Da wir von Anfang an beabsichtigten, echte Prüfungskandidaten zu untersuchen, von einer Simulation im Labor also absahen, griffen wir methodisch gesehen zum Forschungsparadigma des natürlichen Experiments (vgl. McDAVID & HARARI, 1969).

"A natural experiment is essentially only a field study in which Nature and circumstance cooperated by the occurrence of a critical manipulative event. The natural occasion of a certain event may produce a change in conditions which serves the purpose of a manipulated independent variable." (McDAVID & HARARI, 1969, S. 407)

Weil die natürlichen Ereignisse - die Prüfungen - von uns nicht beeinflussbar und manipulierbar waren, werden wir im methodischen Teil nur so weit auf die oben erwähnten Inhalte eingehen, wie sie unserer Planungskompetenz unterlagen.

3.1.1 Selektion der Stichprobe

Aus rein pragmatischen Gründen wählten wir als Untersuchungsstichprobe die Kandidaten eines Prüfungstermins im Hauptfach Psychologie.

Nicht nur die potentielle Kooperationsbereitschaft von Studenten der Psychologie erschien uns größer als die von Studenten anderer Fächer, sondern auch die relative Geübtheit im Umgang mit Fragebögen. Vor allem letzteres erleichtert die zeitliche Planung.

Ein zweiter Aspekt war die potentiell größere Kooperationsbereitschaft des Prüfungsamtes, insbesondere was die Bekanntgabe von Prüfungsterminen und -verschiebungen betrifft.

Ein dritter Grund waren die uns bekannten Räumlichkeiten, was es erleichterte, die Probanden die Fragebögen maximal ungestört ausfüllen zu lassen.

Ein vierter Punkt ist, daß die in Frage kommenden Prüfer (zum Zeitpunkt der Datenerhebung) prinzipiell in beiden Prüfungsarten (Vor- wie Hauptdiplom) Prüfungen abnahmen, so daß wir die Hoffnungen hegen konnten, wenigstens beim Vergleich von Vor- und Hauptdiplomanden Unterschiede nicht auf die Person des Prüfers zurückführen zu müssen.

Theoretisch interessant ist die Aufteilung der Psychologieprüfungen in Vor- und Hauptdiplomprüfungen, die die Operationalisierung des Konzepts der 'Vorerfahrung' sichert. Hauptdiplomanden haben mit Si-

cherheit in ihrem bisherigen Studium mindestens sechs mündliche Prüfungen abgelegt, während Vordiplomanden in ihrer Mehrheit Prüfungsanfänger sein dürften.

Wenn man sich an einem verkürzten Prüfungsbegriff - Prüfung als allgemeinen Leistungsnachweis - orientiert, haben natürlich auch Vordiplomanden in diesem Sinne 'Prüfungserfahrung'. Sie haben im Abitur oder während des zweiten Bildungsweges Prüfungen abgelegt, haben sich Führerscheinprüfungen unterzogen, an sportlichen Wettkämpfen teilgenommen etc. Inwieweit diese Prüfungen die Strategien für mündliche Prüfungen in der Institution Universität beeinflussen, wissen wir nicht. Wir gehen davon aus, daß Hauptdiplomanden mehr Erfahrungen in bezug auf Prüfungen genau der gleichen Institution haben und deshalb die Wahrscheinlichkeit von Kontrollstrategien für Prüfungen in dieser Institution bei ihnen größer ist.

Zum Zeitpunkt der Datenerhebung existierte keine verbindliche Studienordnung, so daß aus der Sicht der Studenten sich die Prüfungsanforderungen nicht aus einer Lehrveranstaltung oder aus einer Sequenz von Kursen und Seminaren ergeben, sondern die im Rahmen der Prüfung anfallenden Probleme in der Prüfung selbst kumulieren.

Die Funktionen von Vor- und Hauptdiplomprüfungen im Studium sind über die Prüfungsordnung hinreichend bestimmbar: Das Vordiplom sichert die Weiterführung des Studiums, sofern in allen Prüfungen eine mindestens 'ausreichende' Leistung erreicht wird, und das Hauptdiplom stellt den berufsqualifizierenden Abschluß dar. Der funktionale Stellenwert der Prüfungen ist also nicht so kompliziert wie in Fächern, in denen die Zensur der Zwischenprüfung den Zugang zu weiterführenden Lehrveranstaltungen

regelt oder die letzte Universitätsprüfung erst einen zweiten Ausbildungsabschnitt **eröffnet wie** bei Juristen und Lehrern.

Als Untersuchungsstichprobe (Kohorte) wählten wir die Kandidaten der Prüfungstermine im Frühjahr 1976. In Anlehnung an MOELLER & SCHEER (1970) nehmen wir die in Abbildung 8 dargestellten Selektionsmechanismen bei der Kohortenbildung an.

Im gestrichelten Kasten im unteren Teil der Abbildung sieht der Leser die untersuchte Kohorte. Wir mußten davon ausgehen, daß nicht alle Kandidaten bereit waren, an der Untersuchung teilzunehmen (= Verweigerer). Ferner mußten wir berücksichtigen, daß einige Kandidaten, die sich für die Prüfung angemeldet hatten und somit in unsere Planung eingingen, ihre Anmeldung noch zurückziehen und somit als Versuchspersonen ausfallen. (Nach der Bochumer Praxis stellt die Anmeldung bis zur Ausschlußfrist eine 'Option zur Prüfungsteilnahme' dar, die ohne Begründung bis unmittelbar vor Beginn der ersten mündlichen Prüfung annulliert werden kann. Da die Ausschlußfrist ca. acht Wochen vor der ersten Prüfung liegt, melden sich einige Kandidaten 'vorsorglich' an, und ein Teil davon zieht die Anmeldung dann später zurück.)

Bei den kooperationswilligen Teilnehmern (= Vordiplomanden und Hauptdiplomanden) unterscheiden wir zwischen denen, die bestanden haben oder durchgefallen sind, denjenigen, die krank wurden und somit für unsere Untersuchung ausfallen (nach Beginn der individuell jeweils ersten Prüfung ist ein Rücktritt

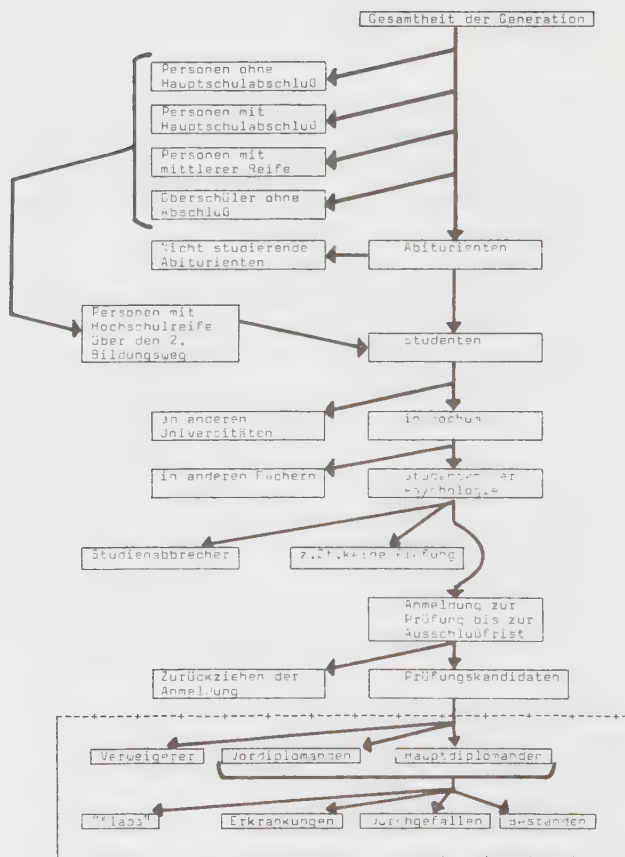


Abbildung 8: Darstellung der Selektionsmechanismen, die zur Bildung der untersuchten Kohorte (im gestrichelten Kasten) beitragen

nur wegen einer ärztlich attestierten Erkrankung möglich) und denjenigen, für die wir den Begriff 'flaps' übernommen haben.

Dieser Begriff bezeichnet Fälle von spektakulären Zusammenbrüchen frühestens eine Woche vor, häufig jedoch während der Prüfung (vgl. HENN, 1951, MALLESON, 1957, 1959). Diese Fälle werden in der Studentenschaft schnell und detailliert kommuniziert und bestimmen die Assoziationen der 'Blockierung' und des 'Ausflippens' zur Prüfungsangst. Über Inzidenz oder Prävalenz finden sich jedoch keine Veröffentlichungen. Trotz der auch uns bekannten anekdotischen mündlichen Berichte nehmen wir deshalb an, daß spektakuläres Versagen ein eher extrem seltenes Phänomen ist, das seine unangemessene Bedeutung allein durch akademischen Klatsch und Tratsch sensationeller Einzelfälle erhält.

3.1.2 Einschränkende Schwierigkeiten der Forschungspraxis

Bei jeder Untersuchung echter Prüfungen werden dem Forscher Einschränkungen auferlegt, deren Begründung nicht forschungsimmanent ist.

Die wichtigste Einschränkung ist die Vertraulichkeit und Nichtöffentlichkeit der Prüfung. Danach ist es nur dem Kandidaten erlaubt, Details über den Prüfungsverlauf an Nichtbeteiligte weiterzugeben. An der Prüfung dürfen nur Prüfer, Beisitzer und Kandidat teilnehmen. (Am Psychologischen In-

stitut Bochum wird in der Praxis von dieser Bestimmung der alten Prüfungsordnung abgewichen und der neuen Prüfungsordnung vorgegriffen: Bei Zustimmung des Kandidaten dürfen andere zuhören. Bei einigen Prüfern besteht eine Tendenz zur Beschränkung der Erlaubnis auf Psychologiestudenten; da jedoch keine Ausweise kontrolliert werden, sind die Prüfungen de facto für jeden, dem der Kandidat zustimmt, zugänglich.)

Die Vertraulichkeit der Prüfungen führt dazu, daß Ton- oder Videoaufnahmen echter Prüfungen - ebenso wie psychophysiologische Ableitungen - keineswegs selbstverständlich sind. In Fällen, in denen dies dennoch geschah (PIEHL, 1972; TANEHI, 1968-1969, pers. Mitteilung 1975; WICK, 1964, zit. n. TURKSON, 1970), handelte es sich vermutlich stets um Prüfungen bei dem unmittelbaren Vorgesetzten oder bei dem direkten Betreuer des Forschers. TIMMERBRINK (1975, pers. Mitteilung) berichtet, daß sich Prüfer von nichtpsychologischen Fächern schlicht weigerten, Tonaufnahmen von mündlichen Prüfungen zu gestatten.

Diese Widrigkeiten schlagen sich in einem allgemeinen Mangel an Studien über Prüfungsverläufe nieder. Der Standard ist die Untersuchung vor oder nach einer Prüfung mit dann retrospektiver Befragung.

Der angeschnittenen Problematik begegneten wir in zwei konkreten Punkten:

1. Der Vorsitzende des Prüfungsausschusses, mit dem wir während der Planungsphase Kontakt auf-

nahmen, meldete Bedenken gegen eine Befragung unmittelbar vor der Prüfung an. Er befürchtete eine Auslösung oder Verstärkung von Ängsten durch die Meßverfahren. Zwar konnten wir auf FISCH (1970) und PIEHL (1972) verweisen, die eine angstreduzierende Funktion von experimentellen Untersuchungen unmittelbar vor einer Prüfung konstatieren, konnten HEINERTH (1972) zitieren, daß klientenzentrierte Gespräche unmittelbar vor einer Prüfung die Angst im gleichen Umfang reduzieren wie Autogenes Training und das Anhören von Musik (Beethoven), und konnten argumentieren, daß wahrscheinlich jedwede Ablenkung unmittelbar vor einer Prüfung angstreduzierend wirkt, konnten jedoch die Nichtreaktivität unserer Meßverfahren nicht nachweisen.

Im Interesse der Prüfungskandidaten verzichteten wir deshalb auf eine Befragung unmittelbar vor Beginn der Prüfung.

2. Einer der Prüfer problematisierte die Darstellung prüferspezifischer Daten. Dies könne zu (gerechtfertigter, vor allem aber ungerechtfertigter) Kritik am individuellen Prüfungsstil führen.

Uns wurde daher nahegelegt, bei der Darstellung der Ergebnisse auch in dieser Hinsicht auf Anonymität zu achten.

Nicht unerhebliche Teile der Forschungsplanung und der Durchführung wurden durch Schwierigkeiten beeinflusst, die nicht methodischer, theoretisch-sachlicher oder technischer Art waren.

3.1.3 Prinzipielle Planung der Datenerhebung

Wir beabsichtigen, die Daten wiederholt zu messen, um die Konstanz der Prüfungsstrategien und den Einfluß einzelner Prüfer wenigstens ansatzweise erfassen zu können.

Zur Erleichterung zukünftiger Prüfungsforschung wollen wir der Vollständigkeit halber den institutionellen Mechanismus am Psychologischen Institut in Bochum zur Zeit der Datenerhebung festhalten.

Für die Prüfungen legt das Prüfungsamt in der Regel einen Prüfungszeitraum von sechs bis acht Wochen fest. Aus diesem Zeitraum wählt der Prüfer Prüfungstage und benennt sie zusammen mit der Anzahl der an diesen Tagen möglichen Prüfungen dem Prüfungsamt. In der Regel favorisieren die Prüfer dabei Zeiträume en bloc, d. h. sie prüfen an möglichst wenigen Tagen hintereinander. Diese Prüfungstage werden vom Prüfungsamt ca. drei Wochen vor Beginn der ersten mündlichen Prüfungen in einer Versammlung aller Kandidaten bekanntgegeben.

Die Kandidaten müssen bei der Anmeldung zur Prüfung, also vor der Ausschußfrist, angeben, welchen Prüfer sie in welchem Fach wählen. Üblich ist, daß zwei Prüfer pro Fach zur Auswahl stehen. In der Versammlung suchen sich die Kandidaten dann geeignete Termine aus, und zwar zuerst die Hauptdiplomanden und dann die Vordiplomanden.

Da jeder einzelne Prüfling bestrebt ist, möglichst große Zeiträume zwischen die einzelnen Termine zu legen, bleiben Konflikte unter den Kandidaten wegen günstiger Termine nicht aus. Nicht selten werden die Konflikte auch dadurch gelöst, daß 'Delegationen' zu den Prüfern geschickt werden und um Terminverschiebung bitten; nach unseren Beobachtungen sind die Bitten nicht selten erfolgreich und werden in der Regel nur dann strikt abgelehnt, wenn der Prüfer am gewünschten Tage auf Reisen ist. Die häufigste Konfliktlösung ist u. E. jedoch, daß diejenigen die günstigen Termine bekommen, die am lautesten schreien (können).

Dieser institutionelle Mechanismus hat für unsere Arbeit zur Folge, daß die tatsächlichen Prüfungstermine uns erst kurz vor Beginn der Datenerhebung bekannt wurden. Wir entschlossen uns daher blind, bei fünf Prüfern, die sowohl von VDN (Vordiploman-

den) als auch HDn (Hauptdiplomanden) gewählt werden konnten, Messungen nach sämtlichen Prüfungen durchzuführen.

Der Flexibilität und der Ökonomie wegen beschränkten wir aus auf die Datenerhebung mittels Fragebögen.

Während der der Datenerhebung vorausgehenden Prüfungsperiode beteiligten wir uns bei allen fünf Prüfern in mindestens fünf Prüfungen als Beisitzer.

Die Teilnahme eines Beisitzers an der Prüfung ist in der Prüfungsordnung zwingend vorgeschrieben, seine Funktion ist jedoch nicht näher definiert. In der Regel handelt es sich um einen Assistenten, der auch das Protokoll führt. Es ist in den Psychologieprüfungen nicht üblich, daß der Beisitzer Fragen stellt, wohl aber ist es üblich, daß Prüfer und Beisitzer nach der Prüfung und vor ihrer Bekanntgabe die Zensur besprechen.

Während unserer 'Vorversuche' ergab sich, daß diese Besprechung mindestens 90 Sekunden dauert; diese Zeitschätzung wurde durch Messungen bei anderen Beisitzern bestätigt. Während dieser mindestens 90 Sekunden wartet der Kandidat im Sekretariat des Prüfers oder auf dem Flur vor dem Prüferzimmer. Geprüft wird in den Arbeitszimmern der Professoren.

Der von uns voraussagbare Prüfungsablauf sah so aus, daß wir bei den fünf Prüfern im Prüfungszeitraum die meisten Kandidaten drei- bis viermal befragen konnten. Ein Prüfling wartet vor der Prüfung in der Regel vor dem Arbeitszimmer des Prüfers, wird dann hereingebeten und nach frühestens 20 Minuten wieder herausgeschickt (von jetzt ab sind Zuhörer nicht mehr zugelassen). Mindestens 90 Sekunden wartet er, während Prüfer und Beisitzer die Zensur besprechen. Dann wird der Prüfling hineinge-

beten und die Zensur vom Prüfer in Anwesenheit des Beisitzers bekanntgegeben.

Nach den Beobachtungen des Autors begründen alle Prüfer die Zensur in mehr als der Hälfte der Fälle. Nach den Prüfungen, in denen mit '1' zensiert wurde, fällt die Begründung meist sehr kurz aus. Regelmäßig erfolgt eine Begründung dann nicht, wenn eine zeitliche Verzögerung eingetreten ist, andere Kandidaten schon warten und die Zensur '2' oder besser ausgefallen ist.

3.1.4 Zusammenfassung und Schlußfolgerung

Zusammenfassend wollen wir die wichtigsten Punkte festhalten:

- Wir betrachten Prüfungen als 'natürliche Ereignisse', wobei Prüfungen mit unterschiedlichen Funktionen sich einmal auf die Prüfungsstrategien, beide zusammen sich dann auf die 'abhängigen' Variablen Prüfungsbelastung und -leistung auswirken.
- Aus pragmatischen Gründen untersuchen wir die Prüfungskandidaten am Bochumer Psychologischen Institut in Prüfungen des Frühjahrstermins 1976.
- Dabei unterscheiden sich Vordiplomanden (VDn) von Hauptdiplomanden (HDn) nach dem Immunisierungsparadigma: HDn hatten mehr Gelegenheit zur Ausbildung von Kontrollstrategien.
- Wir haben es mit einer hochselegierten Kohorte zu tun.
- Wir haben keine Möglichkeit, die natürlichen

Ereignisse zu beeinflussen, und nur eingeschränkte Möglichkeiten für Messungen.

- Wir führen überwiegend retrospektive Messungen durch, was aber dem Standard der bisherigen Forschung entspricht.
- Wir können retrospektiv nach den Prüfungen messen und in den ca. 90 Sekunden, während der Kandidat auf die Bekanntgabe der Zensur wartet.
- In der Planung der Datenerhebung müssen wir uns in einem solchen Ausmaß an den institutionellen Mechanismen orientieren, daß das, was in der klassischen experimentellen Methode 'Design' genannt wird, bei uns zum Ergebnis wird.

Unter diesen methodischen Rahmenbedingungen, in Anbetracht des Hintergrundinteresses der Belastungsprophylaxe und angesichts des derzeit üblichen Methodenstandards müssen wir in besonderem Maße dem Gesichtspunkt des Fehlervarianzrisikos Rechnung tragen. Wir wollen deshalb pauschalen Einwänden entgegentreten, unter 'diesen' Umständen sei sowie so kein 'vernünftiges' Ergebnis erzielbar, bei retrospektiver Befragung sei ein unspezifischer 'Erinnerungseinfluß' viel zu groß und diese 'besondere' Stichprobe (Psychologie - Studenten in Bochum im Frühjahr 1976) sei keineswegs mit anderen Gruppen vergleichbar.

3.2 Maßnahmen, die die Behauptung eines globalen Meßfehlers widerlegen sollen

3.2.1 Zur internen Validität der Datenerhebung: Die Unterscheidung von Peripher- und Kerndaten

Institutionelle Mechanismen oder ethische Überlegungen verboten uns, eine maximale Standardisierung anzustreben. Aus bereits beschriebenen Gründen waren wir auf die Erhebung z. T. retrospektiver Fragebogendaten eingeschränkt.

Wir wählten deshalb für die Datensammlung Daten mit zwei Funktionen. Den ersten Datensatz wollen wir mit 'Kerndaten' bezeichnen; sie dienen der Hypothesenprüfung und -generierung. Den zweiten Datensatz nennen wir 'Peripherdaten'; sie werden lediglich erhoben, um Replikationen früherer Forschungen zu ermöglichen. Die beiden Datensätze sollen im Rahmen dieser Arbeit nicht aufeinander bezogen werden.

Der Entschluß, zwischen zwei Datensätzen zu unterscheiden und beide zu erheben, war Resultat der folgenden Überlegungen.

Unseres Erachtens ist das Fehlervarianzrisiko bei der Methode des natürlichen Experiments besonders hoch, d. h. die interne Validität (vgl. CAMPBELL & STANLEY, 1963, deutsch 1970) ist gefährdet. Weiter oben haben wir beschrieben, daß wir es mit keiner Zufallsstichprobe zu tun haben, daß wir mit

einer Einbuße von Versuchspersonen rechnen müssen, daß die Datenerhebung nach der Prüfung retrospektiv erfolgen muß und deshalb eventuell von Gedächtnisveränderungen betroffen ist, u.a. All dies sind Faktoren, die die interne Validität gefährden können.

Wie meinen nun, daß pauschale Störquellen bei Messungen in der Regel nicht isoliert bei einer einzelnen abhängigen Variablen wirksam sind, sondern diffuser und allgemeiner wirken. Wenn wir in derselben Situation und bei den gleichen Versuchspersonen zwei Datensätze erheben, dann sollten sich pauschale Störquellen auf beide Datensätze auswirken. Dann können wir folgern, daß der Nachweis einer präzisen Messung bei dem einen Datensatz die Annahme einer präzisen Messung beim anderen Datensatz zuläßt.

Den Nachweis präziser Messungen wollen wir über die Peripherdaten antreten. Formal heißt dies, daß wir zunächst über den Nachweis der hohen (formalen) Reliabilität der Peripherdaten auf hohe (formale) Reliabilität der Kerndaten schließen, die in der gleichen Situation von den gleichen Versuchspersonen geliefert werden. Dabei definieren wir hier Reliabilität als Verhältnis von 'wahrer' Varianz zu 'Fehler'-Varianz. Weiter schließen wir dann, daß die Validität der Kerndaten nicht durch niedrige Reliabilität beeinträchtigt sein kann.

Wenn es uns gelingt, schon mehrfach replizierte Ergebnisse in unserer Erhebung erneut zu replizieren (Peripherdaten), dann können wir davon ausgehen, daß unsere Kandidaten mit hinreichender Präzision gemessen wurden und können uns mit ihnen inhaltlich auseinandersetzen.

3.2.2 Die Instrumente zur Erhebung der Peripherdaten und die peripheren Hypothesen

3.2.2.1 Die Freiburger Beschwerdeliste (FBL)

Ein erster Vergleich betrifft den Gesundheitszustand der Kandidaten. Es soll belegt werden, daß die von uns untersuchten Prüflinge nicht kränker oder gesünder sind als andere Angehörige ihres Alters.

Wir nehmen an, daß eine erhöhte Prävalenz oder Inzidenz 'somatischer' Krankheiten kein Resultat z. B. eines erhöhten Infektrisikos ist, daß die Kerndaten nicht durch pathophysiologische Prozesse tangiert werden, sondern daß der ärztlich oder gar amtsärztlich legalisierte Wechsel in die Krankenrolle in verstärktem Maße nur dann auftritt, wenn die Prüfungsordnung dem Kandidaten keine andere Möglichkeit bietet, sich den Anforderungen zu entziehen.

Wir nehmen daher an, daß bei einer Untersuchung ca. zwei bis drei Wochen vor Prüfungsbeginn sich die untersuchte Kohorte nicht von vergleichbaren Stichproben unterscheidet.

Als Methode wählten wir die Freiburger Beschwerdeliste (= FBL) von FAHRENBURG (1976), ein Inventar von 78 somatopsychischen Beschwerden, die danach skaliert werden, ob sie "fast täglich - etwa 3 x Woche - etwa 2 x Monat - etwa 2 x Jahr - nie" subjektiv registriert werden. Die Items werden zu

zehn Symptombereichen zusammengefaßt; eine elfte Skala wird aus der Beschwerdensumme gebildet.

FAHRENBURG berichtet, daß die zehn Symptombereiche den zehn Faktoren einer varimaxrotierten Hauptkomponentenlösung bei 49 % extrahierter Varianz entsprechen. Die Halbierungskoeffizienten betragen zwischen $r_{tt}=.77$ und $r_{tt}=.91$, die Konsistenzkoeffizienten liegen zwischen $\alpha = .70$ und $\alpha = .87$. Alle Skalen korrelieren signifikant mit der Anzahl der Krankenhausaufenthalte, mit der Anzahl der Operationen und den Aufgaben zur Häufigkeit ärztlicher Behandlungen, so daß zumindest eine pauschale Kriteriumsvalidität angenommen werden kann.

Die Liste wurde von uns graphisch den anderen Instrumenten angepaßt, insbesondere verwendeten wir Kurzfassungen der Items statt der Originalfragen (z. B.: Original: "12. Kommen Ihnen in bestimmten Situationen die Tränen?", hier: "Tränen in bestimmten Situationen"; Original: "15. Bleibt Ihnen in aufregenden Situationen die Luft weg, so daß Sie erst wieder ganz tief Atem holen müssen?", in dieser Arbeit: "Atembeschwerden in aufregenden Situationen" usw.)

Diese Maßnahme bringt nicht nur den Vorteil der graphisch-optischen Anpassung an unsere Instrumente, sondern erleichtert auch die Vergleiche mit anderen Prüfungsangstuntersuchungen, die ebenfalls Item-Kurzfassungen verwendeten (TURKSON, 1970; SCHEER & ZENZ, 1973).

Leider entzogen sich sieben Items einer Kurzformulierung, bei der Denotation und Konnotation erhalten bleiben. Diese Items wurden von uns nicht verwendet (Items 10, 11, 20, 22, 44, 47, 51), so daß der von uns verwendete Fragebogen 71 Items enthielt.

3.2.2.2 Das Studiengewohnheitsinventar (SGI) zur Analyse von Arbeitsstörungen

Arbeitsstörungen werden in einen mehr oder weniger direkten Bezug zu Prüfungsängsten gesetzt (vgl. GRAEFE & JÜHRENS, 1971; MOELLER & SCHEER, 1970, 1974; ZIOLKO, 1970; NAEF, 1971; SUAREZ, 1971). Die empirische Forschung beschränkt sich bislang auf die Analyse des Arbeitsverhaltens bzw. inadäquater Arbeitsstile.

MOELLER & SCHEER (1970) untersuchten 162 'normale' Studenten u. a. mit fünfzehn Items eines selbst-konstruierten Arbeitsstilfragebogens. Eine Faktorenanalyse vom Typ R ergab bei 63 % extrahierter Varianz 3 varimaxrotierte Faktoren:

- Sicherheit bzw. Unsicherheit in der Arbeit
(in der Einschätzung des Arbeitserfolges sicher, seltene Zweifel am Arbeitsstil, keine Verlegenheit bei Leistungen vor den Augen anderer, gutes Gedächtnis, keine Kritikempfindlichkeit)
- Große bzw. geringe Besetzung und Stetigkeit in der Arbeit

(regelmäßige Arbeit, wenig Unterbrechungen, eher ehrgeiziger als andere, wenig ermüdbar, häufig Freude an der Arbeit),

- Emotionale Unabhängigkeit bzw. Abhängigkeit in der Arbeit

(Leistung nicht abhängig von Einstellung zu zuteilender Person, Leistung unabhängig von Einschätzung durch andere, häufig Freude an der Arbeit, Bevorzugung selbständiger Arbeit).

Unsicherheit und Abhängigkeit gehen dabei mit depressiven Merkmalen einher (MOELLER & SCHEER, 1970, S. 51 f.) Klienten einer studentischen Beratungsstelle erwiesen sich dabei als unsicherer, weniger stetig und abhängiger von der Einstellung zu Personen, die ihnen ihre Arbeit zuteilen.

Wir werden versuchen, die drei Dimensionen der Sicherheit, der Stetigkeit und der sozialen Unabhängigkeit zu replizieren.

Zu diesem Zweck stellten wir ein Studiengewohnheitsinventar (SGI) zusammen, indem wir 50 Items den Arbeiten von NAEF (1971), SUAREZ (1971), MOELLER & SCHEER (1970) und TURKSON (1970) entnahmen. Das SGI ist im Anhang wiederzufinden.

Von einer Replikation wollen wir dann sprechen, wenn die oben dargestellten Dimensionen sich als geeignet für die Interpretation unserer Faktoren erweisen.

3.2.2.3 Analogskalen zur Messung von Angstverläufen

Ein weiterer peripherer Datensatz stützt sich auf die Arbeiten von FENZ und EPSTEIN (EPSTEIN, 1967; EPSTEIN & FENZ, 1962, 1965).

Danach haben unerfahrene Fallschirmspringer unmittelbar vor dem Sprung die größte Angst; diese Angst ist größer als die von erfahrenen Fallschirmspringern. Ferner wissen wir, daß erfahrene Fallschirmspringer während des Sprungs - wenn es also kein Zurück mehr gibt - stärkere Angst haben als unerfahrene (vgl. EPSTEIN, 1967), wenngleich sie in beiden Fällen geringer ist als unmittelbar vor dem Sprung. EPSTEIN und FENZ konzeptualisieren dies als Erwartungsangst einerseits, die aber andererseits realitätsorientierten Hemmungsprozessen unterworfen ist.

EPSTEINs theoretisches Modell wurde - mit ähnlicher Methodik - von PIEHL (1972) auf Prüfungskandidaten übertragen und verifiziert.

Bei Messungen an Prüfungskandidaten konnte dieser Angstverlauf durch andere Untersuchungen bestätigt werden. MENSEN (1971, zit. n. RACHMAN & BERGOLD, 1976) fand, daß unmittelbar vor der Prüfung die Angst - gemessen mit dem STAI (State-Trait-Anxiety-Inventory nach SPIELBERGER, GORSUCH & LUSHENE, 1970) - ein Maximum erreicht und danach kontinuierlich abfällt. TURKSON (1970) und SCHEER & ZENZ (1973) berichten über ähnliche Verläufe skaliert physiologischer Symptome. WICK (1964, zit. n. TURKSON,

1970), BRIDGES (1973), BROWN & VAN GELDER (1938), PIEHL (1972), FISCH (1970, 1971) berichten übereinstimmend von erhöhten physiologischen Reaktionen während der Prüfungen, die bei einer Nachmessung - mit unterschiedlicher Latenz bei den genannten Autoren - auf 'normale' Werte zurückgingen.

Eines der bestgesicherten Ergebnisse der Prüfungsforschung ist demnach der Angstverlauf. Die Angst ist maximal unmittelbar vor der Prüfung, nimmt dann während der Prüfung etwas ab und geht nach der Prüfung gegen Null. Mit EPSTEIN und FENZ sind zudem Unterschiede zwischen Vor- und Hauptdiplomanden zu erwarten.

Nach FAHRENBURG & WALSCHBURGER (1976), LEVI (1972) und BRIDGES (1973) sind einfache Analogskalen zur Erfassung von Emotionen, die nach dem Muster 'überhaupt keine (EMOTION) ... die stärkste (EMOTION), die ich je hatte' konstruiert werden, den LIKERT-Skalen (Itemkombinationen) nicht unterlegen. Dies gilt für die Erfassung subjektiver Befindensänderung.

Der wichtigste Vorteil von Analogskalen ist ihre Ökonomie. Dies veranlaßte uns, sie für die Messung der emotionalen Zustände vor, während und nach der Prüfung einzusetzen.

Wir verwendeten neunstufige Skalen zur Messung von Angst, Wut, Freude, Depression, Lampenfieber, Hilflosigkeit und Ärger, Aus technischen Gründen konnte die Freude-Skala bei der allerersten Datenerhebung nicht eingesetzt werden. Die Messungen für den

Zeitpunkt 'nach der Prüfung' werden vorgenommen, während der Kandidat auf die Bekanntgabe der Zensur wartet, verlangen also keine Gedächtnisleistungen und werden nicht durch die sichere Kenntnis der Zensur beeinflusst. Die Messungen für die Zeitpunkte 'vor' und 'während der Prüfung' werden retrospektiv nach Bekanntgabe der Zensur durchgeführt.

Die Ergebnisse der Analogskalen haben somit nicht nur für die Beurteilung der globalen Reliabilitätsproblematik Bedeutung, sondern bieten zugleich einen Maßstab für die Verzerrungen, die evtl. bei retrospektiven Datenerhebungen auftreten können.

Wir hypostasieren für den peripheren Datensatz der Analogskalen, daß die Verläufe der Skalen Angst, Lampenfieber und Hilflosigkeit, evtl. auch Depression, mit o. a. Muster folgen, während die Skalen Wut, Ärger und Freude - als Messungen von Emotionen, die zur Gegenkonditionierung von Angstreaktionen verwendet werden können - antagonistisch verlaufen.

3.2.3 Zusammenfassung

Die Peripherdaten werden erhoben, um pauschalen Einwänden, die Fehlervarianz unserer Daten sei generell zu hoch, den Boden zu entziehen.

Ein pauschaler Einwand ist logisch nicht aufrechtzu erhalten, wenn die folgenden Kriterien erfüllt sind:

- Studenten der von uns untersuchten Stichprobe unterscheiden sich nicht von den Eichstichproben FAHRENBERGS (1976).
- In einer Faktorenanalyse über Items zum Arbeitsverhalten lassen sich drei Faktoren mit 'Sicherheit bzw. Unsicherheit in der Arbeit', 'Große bzw. geringe Besetzung und Stetigkeit in der Arbeit' und 'Emotionale Unabhängigkeit bzw. Abhängigkeit in der Arbeit' interpretieren.
- Die Scores der Analogskalen 'Angst', 'Lampenfieber' und 'Hilflosigkeit' fallen kontinuierlich ab, wenn man sie 'vor', 'während' und nach der Prüfung einsetzt.

3.3 Meßinstrumente zur Hypothesenprüfung

3.3.1 Die Messung der emotionalen Belastung

3.3.1.1 Die Messung der Angst:

Das State-Trait-Anxiety-Inventory (STAI)

Intendiert war die Messung aktueller Angst im Sinne der State-Angst (vgl. SPIELBERGER, 1966). Da direkte Messungen während der Prüfungen - z. B. psychophysiologische Ableitungen, Fingerspannungsskalierung (BIRBAUMER et al., 1973), Verhaltensbeobachtung durch Beobachter oder -aufzeichnung durch Videorecorder mit anschließender Beurteilung (GROEGER, 1977) oder Tonaufnahmen mit anschließender

der Analyse von Sprachindices (PIEHL, 1972) - entweder für den Kandidaten unzumutbar und/oder wegen der geltenden Prüfungsordnung verboten und/oder organisatorisch unangemessen aufwendig waren, entschlossen wir uns für den Einsatz eines Fragebogens. Dieser sollte experimentell und klinisch validiert, ökonomisch kurz und hinreichend sensibel für Merkmalsveränderungen sein.

Im Gegensatz zum englischen Sprachraum (vgl. De BONIS, 1974; HARTLAGE, 1972, LADER & MARKS, 1974), wo die Zahl von 'Angst'-Fragebögen nahezu unübersehbar ist, mangelte es (zum Zeitpunkt der Datenerhebung) im deutschen Sprachraum an Tests, die sowohl ökonomisch als auch vom theoretischen Hintergrund her angemessen, zudem an deutschen Stichproben geeicht und validiert sind.

Wir entschlossen uns für die Verwendung der Bonner Übersetzung des STAI (State-Trait-Anxiety-Inventory, SPIELBERGER, GORSUCH & LUSHENE, 1970), Form X-1 (state). Diese Selbstratingskala soll die aktuelle Angst erfassen. Der Test erschien uns besonders geeignet, da er kurz ist (20 Items) und bereits im Zusammenhang mit Examensproblemen eingesetzt wurde (für den englischen Sprachraum: vgl. SPIELBERGER, GORSUCH & LUSHENE, 1970; für den deutschen Sprachraum: MNGEN, 1971, zit. n. RACHMAN & BERGOLD, 1976). Die zu erwartende Kriteriumsvalidität des Tests gewichteten wir höher als die Tatsache, daß zu Beginn der Datenerhebung keinerlei Testkennwerte für die deutsche Version vorlagen.

Für die englischsprachige Version werden Retestkoeffizienten zwischen $r_{tt}=.16$ und $r_{tt}=.54$, Konsistenzkoeffizienten zwischen $\alpha = .83$ und $\alpha = .94$, konkurrente Validitäten zwischen $r_{tt}=.41$ und $r_{tt}=.85$, Korrelationen mit akademischen Leistungsmaßen zwischen $r_{tc} = -.07$ und $r_{tc} = .00$ angegeben (SPIELBERGER, GORSUCH & LUSHENE, 1970).

Die deutsche Version wurde den übrigen von uns verwendeten Tests angepaßt, indem wir die ursprünglich vierstufige Antwortskala ('überhaupt nicht - ein wenig - ziemlich - in starkem Ausmaß') um eine fünfte Stufe ('in sehr starkem Ausmaß') erweiterten. Die von uns verwendete Fragebogenversion ist im Anhang wiedergegeben.

3.3.1.2 Die Messung der Depressivität: Die Selbstrating-Kurzform des Beck- Depression-Inventory (BDI)

Zur Messung depressiver Reaktionen konnten wir zum Zeitpunkt der Versuchsplanung ebenfalls auf kein adäquates deutschsprachiges Verfahren zurückgreifen. Die verfügbaren waren entweder kaum klinisch validiert oder unökonomisch lang.

Wir entschlossen uns deshalb, das BECK-Depressions-Inventar (BDI; BECK et al., 1961, PÖLDINGER et al., 1969) in eine Selbstratingkurzform zu überführen und zusätzlich die Hoffnungslosigkeitsskala (HoS; BECK et al., 1974) zu übersetzen.

Das BDI wurde bislang fast ausschließlich in der psychiatrisch-klinischen Forschung bei Depressiven verwendet und ist ein halbstandardisiertes Verfahren, in dem der Untersucher (Psychiater) im Interview mit dem Patienten für einundzwanzig Kategorien von Symptomen und Verhaltensweisen jeweils vier bis fünf Behauptungen, die einer ordinalskalierten Intensität entsprechen, diskutiert und ankreuzt. Wir entnahmen einundzwanzig Behauptungen mäßiger bis mittlerer Schwierigkeit (Intensität) und ordneten ihnen eine fünfstufige Ratingskala ('stimmt genau 0--0--0--0--0 ganz im Gegenteil') zu. Wir übernahmen die Items und Stufen A1, B2, C1, D1, E1, F1, G0, H1, H2, I1, J1, K1, K3, L1, M1, N1, O1, O1b, R1, U2 und T1. Unberücksichtigt blieben Items aus den Gruppen P, Q und S. Die so entstandene Selbst-Rating-Kurzform ist im Anhang wiedergegeben.

Die von BECK & BEAMESDERFER (1974) vorgelegte englischsprachige Self-Rating-Kurzform stand uns zum Zeitpunkt der Drucklegung der Fragebögen nicht zur Verfügung.

Die von uns verwendete BDI-Kurzform wurde auch von FLIEGEL (1978) eingesetzt. Eine Faktorenanalyse mit $N = 103$ Fragebögen ergab den Eigenwertverlauf der Abbildung 9.

Nach dem Scree-Test (GORSUCH, 1974; GAENSSLEN & SCHUBÖ, 1973) wurde ein Faktor extrahiert und bei 24,95 % extrahierter Varianz die Faktorladungen in Tabelle 2 gefunden.

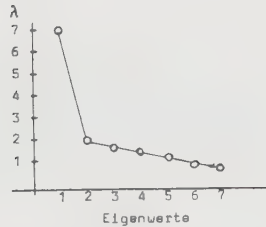


Abbildung 9: Eigenwertverlauf der BDI-Daten von FLIEGEL (1978) mit N = 103

Tabelle 2: Ladungen der Faktorenanalyse und Test-Item-Korrelationen der gewichteten Variablen; Daten von FLIEGEL (1978)

Item	Text	Ladung	$r_{it}(\text{pre})$	$r_{it}(\text{post})$
1.	Es ist mir verleidet, oder: ich bin traurig.	.67	.62	.60
2.	Ich glaube, daß ich nichts zu erhoffen habe.	.57	.51	.48
3.	Ich habe nicht das Gefühl, mehr versagt zu haben als andere Leute.	-.26	.41	.17
4.	Ich habe an nichts mehr dieselbe Freude wie früher.	.74	.62	.72
5.	Ich empfinde mich sehr oft als schlecht oder wertlos.	.73	.71	.75
6.	Ich glaube, daß mir etwas Schlimmes passieren könnte.	.54	.47	.47
7.	Ich bin nicht enttäuscht von mir.	-.44	.55	.49
8.	Ich achte sehr gut auf meine eigenen Irrtümer oder Fehler.	.07		
9.	Ich mache mir Vorwürfe für alles, was nicht klappt.	.55	.57	.56
10.	Ich denke manchmal daran, mir etwas anzutun, aber ich werde es nicht tun.	.6	.51	.72
11.	Ich weine häufiger als gewöhnlich.	.60	.60	.57
12.	Ich bin rascher verärgert (oder gereizt) als gewöhnlich.	.65	.50	.53
13.	Alles, was mich gewöhnlich geärgert hat, berührt mich nicht mehr.	-.25	.77	-.08
14.	Ich interessiere mich weniger als gewöhnlich für andere Leute.	.41	.41	.51
15.	Ich bin jetzt weniger selbstischer und versuche, Entscheidungen aus dem Weg zu gehen.	.71	.63	.61
16.	Ich befürchte, daß ich gealtert oder unvorteilhaft aussehe.	.55	.55	.57
17.	Es braucht eine zusätzliche Anstrengung, an etwas heranzugehen.	.7	.7	.7
18.	Mein Appetit ist nicht so gut wie gewöhnlich.	.55	.55	.54
19.	Ich interessiere mich viel weniger für sexuelle.	.7	.7	.7
20.	Ich bin sehr mit Unwohlsein und Schmerzen beschäftigt, zum Beispiel mit Bauchweh oder Verstopfung oder mit anderen unangenehmen körperlichen Empfindungen.	.71	.43	.7

Item 8 wird wegen der Nullladung ignoriert. Die Items 3 und 13 werden bei der Summenscorebildung negativ gewichtet. Die auf dieser Basis errechneten Korrelationen der gewichteten Items mit dem vorläufigen Testkennwert sind ebenfalls in Tabelle 2 wiedergegeben.

Die Items 13 ('Alles, was mich gewöhnlich gereizt hat, berührt mich nicht mehr') und 19 ('Mein Appetit ist nicht so gut wie gewöhnlich') wurden trotz ihrer Schwäche in der vorläufigen Testform beibehalten.

FLIEGELS Daten ergeben auf der Grundlage der vorläufigen Testform mit 20 Items eine Retestreliabilität (Kontrollgruppe mit $N = 15$) von $r_{tt} = .81$ und eine Split-Half-Reliabilität nach SPEARMAN-BROWN von $r_{1,2} = .90$.

Zum Zeitpunkt des Versandes der ersten Fragebögen waren die eben dargestellten Testkennwerte noch nicht bekannt. Um sicherzugehen übersetzten wir deshalb zusätzlich die Hoffungslosigkeitsskala von BECK et al. (1974) und paßten sie unseren anderen Instrumenten graphisch an. Die Hoffungslosigkeitsskala wurde nur bei der Ersterhebung verwendet, weil sich das BDI als vielversprechend genug erwies. Der Autor geht auf die HoS daher nicht näher ein.

3.3.1.3 Die Messung von Hilflosigkeit:

Die Hilflosigkeitsskala (HiS)

Die Hilflosigkeitsskala (HiS) wurde vom Autor nach der Konzeption SELIGMANS entwickelt (BREITKOPF, 1977).

Die Testentwicklung ging dieser Arbeit zeitlich voraus. Dabei wurden unabhängige Stichproben von Studenten verwandt, die sich in keinem Fall mit der hier untersuchten Kohorte überschneiden.

Die Skala besteht aus zwanzig Items, die alle gleichsinnig gepolt sind. Die bei der Datenerhebung verwendete Version besteht aus zweiundzwanzig Items, da der Computerlauf, demzufolge die Items 7 und 21 ignoriert werden mußten, aus technischen Gründen bei der Drucklegung der Fragebögen noch auf sich warten ließ.

Vor Beginn der Datenerhebung für diese Arbeit war uns bekannt, daß die Split-Half-Reliabilität nach SPEARMAN-BROWN $r_{1,2}=.93$ beträgt und daß nach dem Scree-Test eine einfaktorielle Lösung zu erwarten ist. Die mittlere Korrelation der Items mit dem Gesamtscore beträgt $r_{it}=.77$, und die Daten gaben einen überzeugenden Anlaß, die gewünschte Merkmalsvariation zu vermuten (vgl. BREITKOPF, 1977).

Mittlerweile wurden Retestreliabilitäten von $r_{tt}=.90$ (Kontrollgruppe von GROEGER, 1977) und $r_{tt}=.87$ (Wartekontrollgruppe bei FLIEGEL, 1978) berichtet. FLIEGELS Daten ließen eine Split-Half-Reliabilität von $r_{1,2}=.98$ errechnen (SPEARMAN-

BROWN). Eigene Experimente des Autors ergaben die Unabhängigkeit des HiS-Scores von äußerer Aktivierung, die experimentelle Validierung im triadischen Design, den Nachweis der Populationsunabhängigkeit und zum wiederholten Male den Nachweis der einfaktoriellen Lösung (alles unveröffentlichte Daten).

Für die Summensecorebildung werden für die Stufen O--O--O--O--O die Punkte 5--4--3--2--1 vergeben und von oben nach unten addiert. Die minimale Punktzahl beträgt demnach 20, die maximale 100.

3.3.1.4 Die Skalierung vegetativer Reaktionen: Die Beschwerdeliste

Aus theoretischen Gründen (vgl. SCHACHTER & SINGER, 1962) wie empirischen Befunden (vgl. LACHMAN, 1972, SIMSON, 1971) konnten wir nicht ausschließen, daß die Realisation der Strategien zwar zu vegetativen Reaktionen führt, sich jedoch nicht in den Emotionstests niederschlägt.

Wir entwickelten deshalb einen Kurzfragebogen, indem wir der Freiburger Beschwerdeliste (FAHRENBURG, 1976) aus den Bereichen Allgemeinbefinden, Emotionale Reaktivität, Magen-Darm und Anspannung insgesamt achtzehn Items entnahmen, in Kurzform umformulierten und - analog zur FBL - mit dieser Beschwerdeliste nach der Häufigkeit der Symptome während der Prüfungen fragten. Als Antwortmöglichkeiten waren vier Stufen vorgesehen: 'fast ständig - häufig - 1-2 mal - nie'. Die Prüflinge hatten also

zu skalieren, wie häufig sie die Symptome während der Prüfung registrierten.

Als A-Priori-Auswertung hatten wir die Verteilung der Punktwerte 4-3-2-1 auf die Antwortstufen vorgesehen und die Summensecorebildung durch einfache Addition. Die Beschwerdeliste ist im Anhang wiedergegeben.

3.3.1.5 Analogskalen zur Messung der Post-Prüfungsbelastung

Um uns nicht vollkommen auf retrospektiv erhobene Emotionsmessungen stützen zu müssen, formulierten wir für die Zeit, während der Prüfling auf die Zensur wartete, vier Analogskalen mit jeweils neun Stufen.

Wegen der zeitlichen Begrenzung der Datenerhebungsmöglichkeiten konzentrierten wir uns auf die unmittelbarsten affektiven Komponenten der Post-Prüfungssituation (nach den Ergebnissen von Interviews in der vorausgehenden Prüfungsperiode) und ließen die Kandidaten neunstufige Skalen mit folgenden Fragen und Polen bearbeiten:

1. "Gesetzt den Fall, der Prüfer würde Dir gleich die Chance geben, Deine Zensur durch eine Nachprüfung (5 Min.) zu verbessern. Wie würdest Du reagieren?"

spontan zustimmen (= 1); spontan ablehnen (= 9)

2. "Wie erschöpft fühlst Du Dich in diesem Moment?"

fühle mich voll fit (= 1); so erschöpft fühlte
ich mich noch nie (= 9)

3. "Wie schätzt Du Deine momentane Konzentrationsfähigkeit ein?"

so hoch wie noch nie (= 1); so niedrig wie
noch nie (= 9)

4. "War der Verlauf der Prüfung genau so, wie Du Dir das vorgestellt hattest?"

nein, viel schlechter (= 1); ja, ganz genau so
(= 5); nein, viel besser (= 9)

Die in der Datenerhebung verwendete (graphische) Version kann im Anhang dem Fragebogen AN entnommen werden.

3.3.2 Die Messung der Strategien

3.3.2.1 Strategien I:

Das Testverhaltensdifferential (TVD)

ERTEL (1966) und ERTEL & SCHINDLER (1969 a, b) legten ein Funktionsmodell des Testverhaltens und zu dessen Messung das Testverhaltensdifferential (TVD) vor. Das TVD besteht aus fünfzehn bipolaren Items, von denen je fünf den drei Dimensionen mäßige/erhöhte Anspannung, Variabilität/Stetigkeit und Unifizierung/Diversifizierung zugeordnet sind. Mit der Dimension Anspannung ist dabei nicht z. B.

körperliche Verkrampfung, sondern Aktivität und Initiative beim Problemlösungsverhalten gemeint (vgl. SCHULTE, 1972). Mit Unifizierung ist eine global-synthetisierende Vorgehensweise, mit Diversifizierung eine zergliedernd-analytische beschrieben.

Das TVD wurde im Zusammenhang mit Intelligenztests (ERTEL & SCHINDLER, 1969a, b) und bei der experimentellen Analyse der Feldabhängigkeit (SCHULTE, 1972) verwendet. Die Retestkorrelationen betragen nach SCHULTE (1972) zwischen $r_{tt}=.56$ und $r_{tt}=.87$, sind also zufriedenstellend.

Wir verwendeten das TVD in zwei Formen:

1. in der von ERTEL (1966) vorgestellten Form für die retrospektive Befragung (TVD (HABE) und
2. in einer 'prognostischen' Version, in der der Kandidat skalieren soll, wie er sich in Zukunft zu verhalten gedenkt (TVD (WERDE)).

Die zweite Version entstand aus der ersten durch Transformation des Perfekts in Futur I. Muster der verwendeten Versionen finden sich im Anhang.

In Tabelle 3 haben wir die Items der TVD-Habe-Version in geordneter Reihenfolge wiedergegeben.

Tabelle 3: Die Items des TVD, geordnet
(aus: SCHULTE, 1972, S. 180)

1. <u>Erhöhte Anspannung</u> Bin jetzt besonnen vorgegangen Bin jetzt hartnäckig vorgegangen Bin diesmal gründlich vorgegangen Habe mich diesmal angestrengt Bin diesmal sorgsam vorgegangen	<u>Mäßige Anspannung</u> Bin jetzt zügig vorgegangen Bin jetzt leichtfüßig vorgegangen Bin diesmal rasch vorgegangen Bin diesmal gelassen geblieben Bin diesmal unbekümmert vorgegangen
2. <u>Variabilität</u> Habe jetzt phantasievoll gedacht Habe jetzt die Sache auf mich wirken lassen Bin jetzt ohne strenge Systematik vorgegangen Bin diesmal variabel vorgegangen Habe jetzt Einfälle aufsteigen lassen	<u>Stetigkeit</u> Habe jetzt nüchtern gedacht Habe jetzt die Sache in den Griff genommen Bin jetzt systematisch vorgegangen Bin diesmal stetig vorgegangen Habe jetzt Einfälle unterdrückt
3. <u>Unifizierung</u> Habe jetzt den Gegenstand ganzheitlich aufgefaßt Habe diesmal Zusammenhänge hergestellt Habe jetzt das Verhältnis der Einzelheiten zu einem übergeordneten Ganzen wichtig genommen Habe jetzt den Kern der Sache beachtet Habe jetzt Dinge miteinander verknüpft	<u>Diversifizierung</u> Habe jetzt den Gegenstand zergliedert Habe diesmal die Zusammenhänge aufgelöst Habe jetzt das Verhältnis der Einzelheiten untereinander betont Habe jetzt die vielfältigen Einzelheiten der Sache beachtet Habe jetzt Dinge voneinander getrennt

3.3.2.2 Strategien II:

Die Strategiensammlung (SS)

Neben dem allgemeinen und eher abstrakten Testverhaltensdifferential stellten wir vierund-dreißig Items zu konkreten Prüfungsstrategien zusammen.

Wir selegierten die Items aus den Itempools von ERTEL (1966), TURKSON (1970) und SCHEER & ZENZ (1973). Dabei orientierten wir uns an den Ergebnissen unstrukturierter Interviews mit Kandidaten des vorausgehenden Prüfungstermins, denen wir die Frage stellten, was wichtig sei, welche Strategie richtig sei und was, wenn man es nicht wisse, ganz besondere Verwirrung stifte. Als Resultat ergab sich die Strategiensammlung (SS).

In den Interviews zeigten sich interindividuelle Unterschiede in mehreren Formen. Wir hatten den Eindruck, daß einige Komplexe interindividuelle Unterschiede graduell abbilden. Dies betrifft vor allem die Komplexe der aktiven/passiven und mutigen/vorsichtigen Strategie, wie sie z. B. durch die Items 1, 2, 3, 4, 5 und 6 erfaßt werden. Diese Komplexe stehen in einem nachgewiesenen Zusammenhang mit Prüfungsverhalten.

Andere Komplexe spiegelten kompliziertere, nahezu qualitative interindividuelle Differenzen. Dies schien vor allem Items, die sich mit der Reaktion auf Fragen befassen, zu betreffen, z. B. die Items 11, 14, 17, 20, 22, 29 und 34. Einige interviewte Kandidaten schienen gelassen,

entschieden und differenziert Stellung zu nehmen. Andere äußerten, genau dieses Thema (Item) sei wichtig, sie wüßten aber nicht, welcher Pol richtig sei. Wieder andere gaben sich über- rascht und äußerten, darüber hätten sie noch nicht nachgedacht.

Wir vermuteten daher, daß die Prüfungsstrategien wenigstens zum Teil weniger nach dem Modell der Faktorenanalyse mit linearen, additiven Beziehungen aufgebaut sind, sondern mehr dem Modell der 'Pläne' von MILLER, GALANTER & PRIBRAM (1960, deutsch: 1973) folgen. Prüfungsstrategien entsprechen demnach weniger einem Modell mit wenigen Dimensionen, wobei interindividuelle Unterschiede sich als unterschiedliche 'Stärke' oder 'Intensität' einer oder mehrerer Dimensionen darstellen, sondern mehr einem Modell, in dem hierarchisch vermaschte Regelkreise vorherrschen, wobei interindividuelle Unterschiede sich als Unterschiede im Umfang der einbezogenen Items, in der Sicherheit dieser Einbeziehung und in unterschiedlicher Struktur wiederfinden.

Dies bedeutet, daß wir die vierunddreißig Items der Strategiensammlung in zweifacher Hinsicht auswerten müssen: zum einen in Hinblick auf 'Kernstrategien', zum anderen in Hinblick auf 'Strukturiertheit'.

Unter 'Kernstrategien' verstehen wir Gruppierungen von Items, die in erster Linie graduelle Versuchspersonen-Unterschiede widerspiegeln und auf die theoretische Unterscheidung von 'Gelernt-

ter Hilflosigkeit' und 'Kontrollstrategie' als Pole einer (oder mehrerer) Dimension(en) zu beziehen sind. Als Methode der Wahl betrachten wir hier die Faktorenanalyse. Über den Regressionsansatz werden wir Faktorscores berechnen, die wir dann als Variablen zur Hypothesenprüfung verwenden (vgl. SORGATZ, 1974). Derjenige Faktor oder - multivariat - diejenigen Faktoren, für den (die) wir die von uns postulierten Zusammenhänge und Unterschiede empirisch nachweisen können, beschreibt (beschreiben) die gesuchten Strategien.

Für den Fall (und nur für den Fall), daß sich die angenommene Itemhierarchie so stark auswirken würde, daß sich noch nicht einmal ein Strategiescore über den faktorenanalytischen Ansatz würde ableiten lassen, hatten wir eine Alternativlösung überlegt, die wir im folgenden kurz darstellen wollen:

1. Wir gehen davon aus, daß sich Items, die in einer Itemhierarchie einen geringen Stellenwert haben oder gar nicht zu einer Hierarchie gehören, im Modell der Faktorenanalyse unique Varianzen darstellen. Kernstrategierelevant können demnach nur die Items sein, die bei wiederholter Messung in Faktorenanalysen für jeden Datenerhebungszeitpunkt substantielle Ladungen ($a > |.40|$) erreichen.
2. Eine zweite Möglichkeit der Identifikation von Kernstrategien läuft über Q-Analysen. In Q-Analysen werden die Profile der Versuchspersonen miteinander korreliert, und man erhält als Ergebnis zwei wichtige Informationen:
 - a) über den Scree-Test den Nachweis, daß innerhalb der untersuchten Stichprobe überzufällige Gemeinsamkeiten zu finden sind, und
 - b) mehrere unterscheidbare Gruppen mit gleicher Itemhierarchie.Damit ist zunächst der Nachweis geliefert, daß es - für den Fall inakzeptabler Ergebnisse in den R-Analysen - dennoch Gruppen ähnlicher Antwortmuster gibt.

3. In einem dritten Schritt bestimmen wir diejenigen Items, die zwischen den Gruppen der Q-Analyse Unterschiede wiedergeben. Wegen der (für diesen Fall dann glaubhaften) hierarchischen Itemstruktur kann dies nicht parametrisch multivariat geschehen, sondern univariat auf zwei Wegen:
 - a) Wir rechnen mit den Q-Gruppen als Faktoren univariate Varianzanalysen und bestimmen so ein Set von Variablen, die in allen Datenerhebungszeitpunkten zwischen den dort berechneten Gruppen univariat trennen, und
 - b) wir berechnen Korrelationen zwischen den Q-Faktorladungen und den Items und bestimmen auf diese Weise die maximal typenrelevanten Items. Eventuell ergibt sich für den Punkt b die Möglichkeit, Items, die mit den Ladungen des einen Faktors positiv, mit denen eines anderen Faktors negativ korrelieren, noch hervorzuheben.
4. In einem vierten Schritt schließlich vereinigen wir die verschiedenen Statistikvarianten, indem wir kernstrategierelevante Items über die Vergabe von 'Gewichtungspunkten' identifizieren. Je einen Gewichtungspunkt gibt es für
 1. substantielle Ladungen in allen Meßwiederholungen,
 2. für signifikante F-Werte in den univariaten Varianzanalysen,
 3. für signifikante Korrelationen der Items mit den Q-Faktorladungen, und
 4. für gegenläufige Korrelationen mit Faktorladungen verschiedener Faktoren einer Lösung.

Die so identifizierten Items der 'Kernstrategien' müssen dann noch mit einem Verfahren beurteilt werden, das sowohl hierarchische als auch additiv-lineare Assoziationen erfaßt (auf diesen Punkt gehen wir weiter unten ein).

Für die Analyse der 'Strukturiiertheit' wollen wir zwei unabhängige Ansätze verwenden.

Zum einen haben wir für die vierunddreißig Items bei der Fragebogenkonstruktion nicht nur eine fünfstufige Antwortskala mit den Stufen 'richtig - eher richtig - starkes Problem für mich, weiß ich aber nicht - eher falsch - falsch' vorgesehen, sondern ermöglichen zusätzlich eine Extraantwort 'noch nie darüber nachgedacht'. A priori ist vorgesehen, die Häufigkeit der Extraantworten auszuzählen und so einen pauschalen Strukturiertheitsindex zu gewinnen. LH-theoretisch entspricht dieser Ansatz der (negativen) Bestimmung des Suchraums des problemlösenden Kandidaten: je weniger Items in die Problemerkalkulation aufgenommen werden, desto kleiner sollte der Suchraum ausfallen.

Zum anderen wollen wir diejenigen Items, die von der Mehrheit der Kandidaten gleichsinnig beantwortet werden, identifizieren. Diese Items beschreiben die 'Selbstverständlichkeiten' und stellen eine Reflexion des Prüfungssystems dar. Wir wollen diese Items 'Trivialitems' nennen, weil aus der - gleichsinnigen - Reaktion der Gesamtgruppe geschlossen werden muß, daß sie 'Triviale' erfassen. LH-theoretisch stellen sie diskriminative Reize dar, und wir meinen, daß diejenigen Kandidaten, die sich bei den Trivialitems 'differenzierte' Reaktionen leisten, ihren Strategiesuchraum unnötig ausweiten bzw. die Diskrimination verschiedener Bedingungen nicht geleistet haben (vgl. Kap. 2.2.2.3.1).

Kernstrategien, Trivialitems und pauschaler Strukturiertheitsindex stellen drei verschiedene methodische Varianten bei der Identifikation von Strategien dar.

Einzelheiten stellen wir im Ergebnisteil dar.

Die von uns in der Datenerhebung verwendete Strategiensammlung ist im Anhang wiedergegeben.

3.3.3 Zur Validierung durch Fremdbeobachter: Das Beisitzerblatt

Zum Zeitpunkt der Planung der Datenerhebung erschien es uns günstig, den objektiven Prüfungsverlauf wenigstens ansatzweise und so zu erfassen, daß gegen das Vertraulichkeitsprinzip in Prüfungen nicht verstoßen wird.

Der wichtigste Punkt ist die Reliabilität der Zensur, die wir formal als übereinstimmende Leistungsbeurteilung durch Prüfer und Beisitzer konzipierten. Wir entwarfen ein Beisitzerblatt und stellten darauf die Frage: "Weicht Ihre Einschätzung der Prüfungsleistung des Kandidaten von der endgültigen Zensur ab? (= Ihre Einschätzung vor der Diskussion mit dem Prüfer)".

Weiterhin interessierte uns die Validität der Emotionstests. Zu diesem Zweck ließen wir die Beisitzer auf einer neunstufigen Analogskala mit den Polen 'extrem ruhig' (= 1) und 'extrem aufgeregt' (= 9) die Aufregung des Kandidaten a) zu Beginn und b) am Ende der Prüfung skalieren.

Drittens interessierte uns die Prüfungssteuerung durch den Prüfer. Wegen der für das Ausfüllen des

Beisitzerblattes sehr knapp bemessenen Zeit beschränkten wir uns auf zwei Items.

1. "Hat der Prüfer den Kandidaten immer ausreden lassen? ... Wenn nicht, wie häufig hat der Prüfer den Kandidaten unterbrochen?"
2. "Konnten Sie beim Prüfer während der Prüfung Gefühle beobachten?"

Zu 1.: Eine Unterbrechung durch den Prüfer bedeutet, daß dem Kandidaten explizit die Steuerung der Interaktion entzogen und vom Prüfer wahrgenommen wird.

Zu 2.: Sichtbarer Ausdruck von Gefühlen durch einen Prüfer verstößt gegen die von uns angenommene professionelle Prüferrolle.

Viertens interessierte uns die Interaktionssteuerung durch den Kandidaten. Wegen der knapp bemessenen Zeit beschränkten wir uns auch hier; als Reaktionsstichprobe für die Fremdbeurteilung wählten wir die Beantwortung von Fragen. Kann ein Kandidat nämlich eine Frage in einer Prüfung nicht beantworten, dann bleiben dem Prüfling - nach unseren Beobachtungen in der vorausgehenden Prüfungssaison - lediglich drei Möglichkeiten:

1. Er kann die Frage zurückweisen bzw. seine Unkenntnis zugeben, d. h. er kann eine konkrete Steuerungsreaktion des Prüfers blockieren, löschen,
2. er kann versuchen, über seine Gefühle zu reden, um so den Prüfer zu einer 'Mitleidskorrektur' zu veranlassen, und
3. er kann 'drum herum reden', d. h. er kann auf eine Frage antworten, die ihm gestellt worden sein könnte mit der Absicht, den Prüfer zu besänftigen oder ihn einzulullen oder ihn zu veranlassen, weitere Informationen zu geben.

Wir ließen die Beisitzer die Häufigkeit dieser Varianten skalieren.

Ein Muster des Beisitzerblatts befindet sich im Anhang.

Da wir nicht davon ausgehen konnten, von den Beisitzern als untrainierten Beobachtern valide und reliable Ratings zu erhalten (die Ausbildung zum Diplompsychologen beinhaltet nicht notwendig ein Beobachtertraining), formulierten wir die Fragen möglichst verhaltensnah und überprüften sie zusätzlich, indem wir zufällig anwesende Prüfungszuschauer ebenfalls den Beisitzerfragebogen ausfüllen ließen. Das Übereinstimmungsmaß liefert uns einen Hinweis, inwieweit der Einsatz zufälliger oder untrainierter Beobachter überhaupt valide Daten liefern kann.

3.3.4 Zum Zielzustand des Kandidaten:

Die Leistungsdaten

Wir haben die mündliche Prüfung aus der Sicht des Kandidaten als Problem mit drei Bestimmungsmerkmalen definiert: der Ausgangssituation, den verknüpfenden Operationen und dem Zielzustand. Als angestrebten Zielzustand hatten wir eine möglichst gute Zensur definiert.

Bei der Baseline-Datenerhebung (14 Tage vor der ersten mündlichen Prüfung) stellten wir den Kandidaten die Fragen:

- "Welche Zensur erwartest Du als Gesamtzensur?" und
- "Welche Zensur erwartest Du in der Mehrheit der mündlichen Einzelprüfungen?"

Für beide Fragen gaben wir vier Antwortkategorien vor ('eher 1', 'eher 2', 'eher 3', 'eher 4').

Ferner entwickelten wir zwei neunstufige Analogskalen, um die prospektiven emotionalen Reaktionen auf die Gesamtzensur erfassen zu können:

- "Wie wären Deine Gefühle bei einer Gesamtzensur '3'?"

(Freude = 1, Gleichgültigkeit = 5, starker Ärger = 9)

- "Wie wären Deine Gefühle bei einer Gesamtzensur '1'?"

(extreme Freude = 1, völlige Verblüffung = 9).

Für die Zeit, während der Kandidat auf die Bekanntgabe der Zensur wartet, sahen wir zwei Fragen vor:

- "Welche Zensur wirst Du bekommen?"
- "Mit welchem Prozentsatz des geprüften Stoffes hast Du Dich bereits während des Studiums (= frühere Semester) beschäftigt?"
(0, 10, 25, 50, 75, 90, 100 %)

Die tatsächlich erhaltene Zensur eruierten wir in der ersten Frage, die wir dem Kandidaten nach ihrer Bekanntgabe stellten.

3.3.5 Zusammenfassung

Die emotionalen Belastungen in der Prüfung erfassen wir retrospektiv mit dem STAI, dem BDI, der HiS und der Beschwerdeliste. Nach der Prüfung messen wir - während der Kandidat auf die Bekanntgabe der Zensur wartet - die Postprüfungsbelastung on-line mit vier Analogskalen. Fremdbeurteiler (Beisitzer und Zuschauer) skalieren die Aufregung zu Beginn und gegen Ende der Prüfung. Für die Hypothesenprüfung sollen in erster Linie die drei Emotionstests HiS, STAI und BDI eingesetzt werden.

Die Prüfungsstrategien messen wir zum einen retrospektiv mit dem TVD(Habe) und über die fremdskalierten (durch Beisitzer und Zuschauer) Items der Prüfungssteuerung durch den Prüfer einerseits und

den Kandidaten andererseits. Zum anderen messen wir die Prüfungsstrategien prospektiv mit dem TVD(Werde) und der Strategiensammlung.

Im Mittelpunkt der Leistungsdaten steht die erhaltene Zensur, die wir on-line nach der Bekanntgabe eruieren. Weiter erheben wir on-line die Zensur, die der Kandidat während des Wartens erwartet und (zu Beginn der Datenerhebung) die erwartete Gesamtzensur. Retrospektiv erheben wir nach der Prüfung den bereits während des Studiums bekannt gewordenen und dann geprüften Stoff (in Prozent) und prospektiv die emotionale Reaktion auf eine Gesamtzensur '1' und auf '3' (bei der Baselinedatenerhebung). Als Reliabilitätskriterium von der Fremdbeurteilung her wählen wir die abweichende Zensureinschätzung durch den Beisitzer.

3.4 Datensätze zur Hypothesengenerierung

Wir haben es für notwendig gehalten, bei den Kern-
daten zwischen Datensätzen zu unterscheiden, die wir zur Hypothesenprüfung benötigen, und Datensätzen, deren Struktur wir nicht kennen, die aber zur theoretischen Weiterentwicklung, zur praktischen Arbeit oder zur Erweiterung des Explanandums von Interesse sind.

Die im folgenden beschriebenen Datensätze unterscheiden sich von den Peripherdaten einmal dadurch, daß sie erstmals in der Prüfungsforschung erhoben werden, zum zweiten dadurch, daß spezifizierte

Hypothesen nur a posteriori möglich sind. Dies bedeutet nicht, daß wir keine globalen Vorstellungen von potentiellen Ergebnissen hätten, sondern daß wir diese Vorstellungen nicht durch Rekurs auf Referenzarbeiten spezifizieren können, oder aber - wie im Falle der Kausalattribution - die Zusammenschau der Referenzarbeiten ein verwirrendes Bild ergibt.

3.4.1 Zur Wahl der Prüfer durch den Kandidaten

Ein Gesichtspunkt, den wir in diesem Kapitel diskutieren wollen, ist die Frage nach der Auswahl der Prüfer.

Psychologiestudenten können in der Regel zwischen mindestens zwei Prüfern in jedem Fach wählen. Die Kandidaten treffen bei der Anmeldung zur Prüfung eine verbindliche Wahl. Es gibt bislang keinerlei Untersuchungen, die abklären, welche Variablen zur Entscheidungsfindung beitragen.

Im Rahmen dieser Arbeit ist diese Fragestellung insofern relevant, als der Kandidat zu einem sehr frühen Zeitpunkt eine strategische, irreversible Entscheidung treffen muß und sich für den einen oder gegen den anderen Prüfer festzulegen hat.

Die plausibelste 'Hypothese' ist, daß die Kandidaten den Weg des geringsten Widerstandes gehen und sich jeweils für den Prüfer entscheiden, dessen Prüfung sie für die leichteste halten. Diese logische und ökonomische Entscheidungskalkulation muß

jedoch nicht zutreffen, da es interindividuelle Unterschiede in der Schwierigkeitspräferenz gibt (vgl. HECKHAUSEN, 1976), die möglicherweise in die Bevorzugung von Prüfern einfließen. Ferner wissen wir, daß Arbeitsgruppen dazu neigen, geschlossen einen Prüfer zu wählen, was nicht unbedingt ein Abbild der individuellen Präferenz sein muß.

Wir befragten die Kandidaten vor Beginn der Prüfungsserie nach ihrer Angst vor und ihrer Einschätzung der Prüfungsanforderungen bei den beteiligten Prüfern, zusätzlich bei je einem Prüfer, bei dem Vor- bzw. Hauptdiplomanden sich nicht prüfen lassen können. Die Originalfragen können dem Fragebogen 'Prüfungsstress' im Anhang entnommen werden.

Somit können wir Kontraste zwischen wählbaren Prüfern bilden. Unsere 'Hypothese' lautet, daß von den wählbaren Prüfern derjenige tatsächlich gewählt wird, der weniger Angst auslöst und dessen Anforderungen geringer skaliert werden. Ferner können wir durch Vergleich der wählbaren mit den nicht wählbaren Prüfern abklären, inwieweit die tatsächliche Entscheidungsmöglichkeit die Skalierung von Angst und Prüfungsanforderungen beeinflußt.

Diese 'Hypothese' widerspricht nicht unserer im theoretischen Teil vorgetragenen Auffassung, daß es beim Verhalten in der Prüfung zwischen den Prüfern keine ins Gewicht fallenden Unterschiede gibt ('Prüfer prüfen profihaft').

3.4.2 Die Kausalattribution des Prüfungsergebnisses

Als Ergebnis, dessen Kausalattribution es zu erfassen galt, definierten wir die Zensur.

Zur Analyse der Attribution formulierten wir für die vier Pole der zwei Attributionsdimensionen je drei Items:

Tabelle 4: Liste der Attributionselemente

<u>Attributionselement</u>	<u>Item</u>
internal-variabel	1. Ich habe mich angestrengt. 5. Ich war faul. 9. Ich war fleißig.
external-variabel	2. Ich hatte Glück. 6. Ich hatte Pech. 10. Das war Zufall.
internal-konstant	3. Das liegt an meinen Fähigkeiten. 7. Das liegt an meiner Unfähigkeit. 11. Ich bin intelligent.
external-konstant	4. Der Prüfer prüft gut. 8. Der Prüfer prüft schlecht. 12. Es wird nicht viel verlangt.

Als dazugehörige Instruktion formulierten wir: "Worauf führst Du Dein Abschneiden zurück? In welchem Ausmaß siehst Du die Gründe dafür in den folgenden Ursachen?"

Als Antwortmöglichkeit war eine fünfstufige Skala mit den Stufen 'bestimmt nicht - vermutlich nicht - vielleicht doch - ziemlich sicher - völlig sicher' vorgesehen.

Die A-Priori-Auswertung sah vor, für die Antwortstufen die Punkte -2/-1/+1/+2/+3 zu vergeben und für jeden Pol einen Kennwert zu bilden, indem das erste und dritte Items addiert und das zweite Item, das negativ formuliert war, subtrahiert wird. Der Kennwert für den internal-variablen Pol IV wurde demnach z. B. so berechnet:

$$\begin{aligned} \text{IV} &= \text{Zustimmung Anstrengung plus Zustimmung} \\ &\quad \text{Fleiß minus Ablehnung Faulheit} \\ &\quad (+2) + (+3) - (-1) = 6 \quad (\text{Beispiel 1}) \end{aligned}$$

oder

$$\begin{aligned} \text{IV} &= \text{Ablehnung Anstrengung plus Ablehnung} \\ &\quad \text{Fleiß minus Zustimmung Faulheit} \\ &\quad (-2) + (-1) - (+3) = -6 \quad (\text{Beispiel 2}) \end{aligned}$$

In der Datenerhebung vierzehn Tage vor der ersten mündlichen Prüfung setzten wir die Attribuierungsskalen zweimal ein. Die Versionen unterschieden sich durch die Frage:

1. "Gesetzt den Fall, Du würdest in einer Prüfung besser abschneiden, als Du selbst es erwartet hast. In welchem Ausmaß würdest Du die Gründe dafür in den folgenden Ursachen sehen?" und
2. "Gesetzt den Fall, Du würdest in einer Prüfung schlechter abschneiden, als Du selbst es erwartet hast. In welchem Ausmaß würdest Du die Gründe dafür in den folgenden Ursachen sehen?"

3.4.3 Die emotionale Reaktion auf die Zensur - die Zensurkonsequenzen

Die emotionalen Reaktionen auf Prüfungszensuren sind bislang nicht untersucht worden.

Unmittelbar im Anschluß an die Frage im Fragebogen F2 nach der erhaltenen Zensur stellten wir die Frage: "Welche Gefühle löst die Zensur in Dir aus? (Bitte alles Zutreffende ankreuzen!)".

Als Antwortmöglichkeit sahen wir zwölf Items vor, die wir im folgenden gruppiert wiedergeben (die in der Datenerhebung verwendete Version ist im Anhang dem Fragebogen F2 zu entnehmen):

- Erleichterung, Enttäuschung,
- Ärger, Wut,
- Freude, Zufriedenheit, Zuversicht, Sicherheit
- Angst, depressive Gefühle, Verwirrung, Hilflosigkeit.

3.4.4 Die Variablen zur Erhebung der Prüfungs- anmeldemotivation

Wir haben angenommen, daß die unterschiedliche Funktionalität von Prüfungen im Rahmen des Prüfungssystems sich psychisch widerspiegeln müßte (Hypothesengenerierungskomplex B).

Um dies festzustellen, interviewten wir in dem der Datenerhebung vorausgehenden Prüfungszeitraum Prüfungskandidaten hinsichtlich ihrer Gründe, sich für die Prüfung anzumelden, und hinsichtlich ihrer

Einstellung für den Fall des Falles, sie würden durchfallen. Die erhaltenen Antworten ließen sich (ohne Gewichtung) zu zehn (Anmeldung) bzw. zehn (Nicht-Bestehen der Prüfung) Themen zusammenfassen.

Wir wiederholten in der Datenerhebung die **Fragen** und formulierten für jedes Thema eine vorgegebene Antwort: "Warum hast Du Dich für die Prüfungen gemeldet? (Bitte alles Zutreffende ankreuzen!)"

1. weil man üblicherweise genau in meinem Semester die Prüfungen ablegt,
2. weil ich in den Beruf will,
3. aus finanziellen Gründen,
4. damit ich mit meiner Ausbildung weitermachen kann,
5. weil ich sowieso schon zu lange studiert habe,
6. weil man mir eine Stelle angeboten hat,
7. weil das bisherige Studium langweilig geworden ist,
8. weil meine Angehörigen (Eltern, Freund(in), Frau o. a.) das erwarten,
9. weil ich ausreichend vorbereitet bin,
10. weil ich etwas anderes machen möchte.

"Gesetzt den Fall, Du müßtest eine Prüfung wiederholen, was würde geschehen? (Bitte alles Zutreffende ankreuzen!)"

11. Meine Angehörigen wären ausgesprochen ärgerlich.

12. Ich käme in finanzielle Schwierigkeiten.
13. Ich wäre vor lauter Scham sehr unsicher im Umgang mit anderen.
14. Das würde wegen des erneuten Arbeitsaufwandes für mich die Grenze der Belastbarkeit darstellen.
15. Das würde für mich wirklich nichts anderes bedeuten als ein verplempertes Semester.
16. Ich würde zwar die Wiederholungsprüfung anstreben, wäre aber unsicher, ob ein Weiterstudium noch Zweck hat.
17. Im ersten Moment würde ich die Welt nicht mehr verstehen; darüber würde ich aber hinwegkommen.
18. Das ist eine unbeantwortbare Frage, weil die Wahrscheinlichkeit durchzufallen für mich nur theoretisch besteht.
19. Das würde mich nicht wundern, weil ich für einige Prüfungen sowieso ein Risiko einkalkuliert habe.
20. Das wäre sehr ärgerlich, würde mich aber nicht umwerfen, weil ich es ja eigentlich nicht verhindern konnte.

Die unterstrichenen Worte markieren die in den mündlichen Äußerungen betonten Inhalte. Wie wir Fragen und vorgegebene Antworten in der Datenerhebung verwendeten, ist aus dem Fragebogenteil 'Prüfungsstress' im Anhang zu entnehmen.

3.5 Datenerhebungs- und Datenverarbeitungsplan

3.5.1 Konkrete Planung der Datenerhebung

Von uns wurde vorgesehen, ca. drei Wochen vor Beginn der mündlichen Prüfungen einen Fragebogen mit der Post an alle Prüfungskandidaten zu versenden.

In diesem Fragebogen wird u. a. die Baseline für die Emotionstests erhoben; wir kodieren die damit erhaltenen Daten mit dem Datenerhebungszeitpunkt-kürzel 'BL'.

Der Fragebogen besteht aus acht Fragen zur Person, acht Items zur Globaleinschätzung von Prüfern und Zensuren, acht Fragen zur Vorerfahrung mit Prüfungen, zwei Fragenkomplexe zum Motiv der Prüfungsanmeldung, zwei Attribulierungsskalen, den Items zur Baselineerhebung des TVD, der HiS, des BDI, des STAI, der SS, ferner die HoS und für die Erhebung der Peripherdaten das SGI, die FBL und sechs (von sieben) Analogskalen.

Nach der mündlichen Prüfung sollte jeder Kandidat, während er auf die Zensur wartet, einen kurzen Fragebogen (AN) mit Fragen zur erwarteten Zensur, zum Prüfungsverlauf und -inhalt, zur momentanen Erschöpfung und Konzentrationsfähigkeit, zur Reaktion auf eine Nachprüfung sowie (nunmehr) sieben Analogskalen beantworten. Als 'Schockabsorber' dienten zwei Fragen zum Namen des Prüfers und zur Anzahl der bereits abgelegten Prüfungen.

Eindeutiges Ergebnis von Vorversuchen war, daß AN von Studenten binnen 60 Sekunden ausgefüllt werden konnte.

Nach Bekanntgabe der Zensur füllte der Beisitzer sein Fragenblatt und der Prüfling die Fragebögen 'F2' und 'F3' aus. In letzteren wurden die Zensur, die emotionale Reaktion auf die Zensur und die Zensurattribuierung, mit je sieben Analogskalen die emotionalen Reaktionen vor und während der Prüfung, mit dem TVD die Vorgehensweise während (Formulierung 'habe', 'bin', in den Tabellen kodiert als TVD(H) der gerade abgelegten und bei zukünftigen Prüfungen (Formulierung 'werde', in den Tabellen kodiert als TVD(W), sowie die SS für künftige Prüfungen erfaßt. Die emotionalen Reaktionen wurden somit retrospektiv erfaßt.

Aufgrund von Erfahrungen in Vorversuchen konnten wir nicht ausschließen, daß ein größerer Prozentsatz der Kandidaten die Bearbeitung dieses Item-pools (Bearbeitungszeit ca. 15 bis 30 Minuten) schlicht verweigern würde. Wir unterteilten den Pool deshalb in zwei Teile (F2 und F3), wobei in F2 die retrospektiv zu erhebenden Teile aufgenommen wurden, in F3 die prospektiv zu skalierenden Strategien (TVD(W) und SS). Wir planten, bei Unwilligkeit der Versuchspersonen, F3 mit nach Hause mitzugeben und evtl. Verzerrungen in Kauf zu nehmen. Wir nahmen dabei an, daß prospektive Daten den durch Erholung und Verstreichen der Zeit bedingten Verzerrungen weniger unterworfen sein würden als retrospektiv erhobene Daten.

Zur Datenerhebung wurden alle Prüfungen bei fünf Prüfern¹ ausgewählt, die sowohl von Vor- als auch von Hauptdiplomanden gewählt werden konnten. Den Kandidaten wurde bei jeder Prüfung, die sie bei einem der Prüfer ablegten, die Fragebögen übergeben.

Bei der Verteilung der Termine Ende Dezember 1975 wurden die Kandidaten über die Erhebung informiert und zur Teilnahme aufgefordert. Wir ließen Listen herumgehen, in die sich die Teilnahmewilligen mit ihren Adressen eintragen konnten. Wir erhielten 12 Unterschriften von HDn und 37 von VDn. Einen der Listenbögen haben wir dem Anhang hinzugefügt. Ein Vergleich mit den Unterlagen des Prüfungsamtes ergab, daß bei Ausschlußfrist insgesamt 24 HDn und 44 VDn zur Prüfung gemeldet waren. Die uns theoretisch zur Verfügung stehende Population bzw. Stichprobe betrug somit maximal 68 Personen. Wir glaubten, mit 49 Versuchspersonen sicher rechnen zu können. Diese Mengen erschienen uns ausreichend, um die Erhebung trotz des Motalitätsrisikos (CAMPBELL & STANLEY, 1963) riskieren zu können.

Nach der Verteilung der Termine ergab sich der Zeitplan der Abbildung 10 und der Tabelle 5!

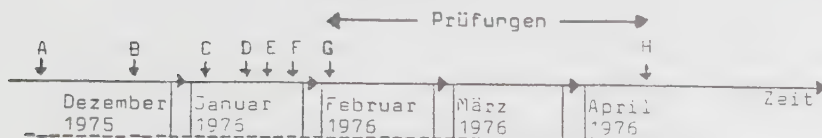


Abbildung 10: Zeitplan für die Datenerhebung;
Erläuterungen s. Tabelle 5!

¹Wir danken an dieser Stelle den beteiligten Professoren für ihr Verständnis und ihre Duldsamkeit.

Tabelle 5: Zeitplan für die Datenerhebung

Datum	Aktivität
A) 07.12.1975	Schreiben an den Vorsitzenden des Prüfungsausschusses mit kurzer Darstellung des Projekts und der Bitte, die Namen der Kandidaten, die Prüfungstermine und etwaige Verschiebungen mitzuteilen; als Folge davon die Teilnahme an der
B) ca. 18.12.75	Verteilung der Prüfungstermine, i. e. Festlegung der Datenerhebungszeitpunkte.
C) 08.01.1976	Schreiben an diejenigen Kandidaten, die sich bei der Verteilung der Termine (B) nicht durch Eintragen in eine Liste zur Teilnahme bereit erklärt hatten.
D) 16.01.1976	Versenden der Fragebögen unter Beilage eines frankierten Briefumschlags an alle (jetzt nur noch) 63 potentiellen Teilnehmer.
E) 23.01.1976	Mahnschreiben an die Kandidaten, die ihren Baselinefragebogen noch nicht zurückgesandt hatten.
F) 30.01.1976	Schreiben an alle in Frage kommenden Beisitzer mit der Bitte um Unterstützung.
G) 02.02.1976	Tag der ersten mündlichen Prüfung für die ersten Kandidaten beim ersten der fünf Prüfer.
H) 05.04.1976	Tag der letzten mündlichen Prüfung für die letzten Kandidaten beim letzten der fünf Prüfer.

Der Vorgehensplan für die Datenerhebung während der Prüfungen sah ein kurzes mündliches Gespräch mit dem Kandidaten vor der Prüfung vor, in dem wir auf die nach der Prüfung folgende Befragung hinwiesen und dem Prüfling alles Gute wünschten. Nach der Prüfung wurden die Fragebögen AN, F2 und F3 persönlich übergeben und der Kandidat auf ein (relativ) ruhiges Plätzchen zum Ausfüllen aufmerksam gemacht. Während des Ausfüllens hielt sich der Autor in der Nähe des arbeitenden Kandidaten auf, vor allem, um eine Störung der Datenerhebung durch neugierige Kommilitonen weitestgehend einzuschränken.

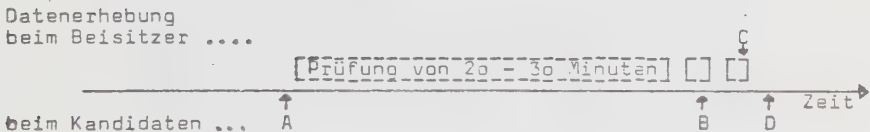


Abbildung 11: Datenerhebungsplan für die einzelnen Prüfungen:

- a) kurzes mündliches Gespräch,
- b) Ausfüllen des Fragebogens AN;
Prüfer und Beisitzer diskutieren
die Zensur,
- c) Ausfüllen des Beisitzerblatts durch
den Beisitzer; der Prüfer teilt dem
Prüfling die Zensur mit,
- d) Ausfüllen der Fragebögen F2 und F3.

3.5.2 Datenschutzmaßnahmen

Wir hatten sowohl den Kandidaten als auch den Prüfern Anonymität zugesichert. Wegen der wiederholten Messungen benötigten wir jedoch eine Identifikations-

möglichkeit; ebenso galt es, die Zuordnung der Beisitzerblätter zu ermöglichen. Wir wählten folgendes Vorgehen:

Den Kandidaten wurde empfohlen, als Code für den Baselinefragebogen, der über die Bundespost zurückkam, ihre Matrikelnummer zu wählen. Dies ist einerseits ein Code, der nicht vergessen werden kann, andererseits wird der Kreis derjenigen, die die Möglichkeit haben, über die Matrikelnummer die Person zu identifizieren, stark eingeschränkt.

Nun wurde bei jedem eingehenden Fragebogen eine laufende Nummer eingesetzt, die Zuordnung Matrikelnummer-laufende Nummer notiert und das Deckblatt des BL-Fragebogens mit der Matrikelnummer vernichtet. Die laufende Nummer blieb während des gesamten Projekts dem Kandidaten zugeordnet. Auf den Fragebögen wurde lediglich die laufende Nummer als Code und eine Kennzahl für die erste, zweite usw. Meßwiederholung notiert.

Zusätzlich legten wir eine Liste mit Namen und Matrikelnummer an. Wählten die Kandidaten bei der BL-Erhebung einen freien Code, so nahmen wir diesen statt der Matrikelnummer auf. Da wir die Fragebögen persönlich übergaben, war der freie Code spätestens nach der zweiten Meßwiederholung für uns kein Geheimnis.

Wir hatten die Möglichkeit, jederzeit mit Hilfe der Namen-Matrikelliste und der Matrikel-laufende - Nummerliste die Fragebögen zu identifizieren; ohne eine dieser Listen war eine Identifikation aber nicht möglich. Nach Abschluß der Datenerhebung wurden beide Listen vernichtet, so daß nirgendwo archiviert ist, wer an unserer Erhebung teilnahm und wie er die Fragebögen ausfüllte.

Die Beisitzer erhielten ihr Blatt mit der laufenden Nummer versehen und einem angehefteten Zettel mit dem Namen des Prüflings. Den Namen entfernten sie nach dem Ausfüllen und gaben die Bögen zurück. Eine Identifikation durch Dritte sollte daher ausgeschlossen sein.

Es ist selbstverständlich, daß für Zugriffe auf die EDV-Dateien auch noch die Kenntnis des Ablochplans notwendig ist. Da bei den Fragebögen die Anonymität der Prüfer nicht gewahrt werden konnte, sorgten wir bei der Ablochung dafür.

Es versteht sich von selbst, daß während der Datenverarbeitung die üblichen computertechnischen Sicherheitsmaßnahmen beachtet wurden.

3.5.3 Darstellung der Konfigurationsfrequenzanalyse und Begründung für ihren Einsatz als Hauptanalysemethode

Die bei der Auswertung verwendeten parametrischen Statistikprogramme sind bereits so häufig einschließ-lich ihrer mathematischen Grundlagen dargestellt worden, daß wir auf eine Darstellung hier verzichten und auf GORSUCH (1974), BEUTEL et al. (1976), HAKSTIAN & KYUNG (1973), FINN (1972) und GAENSSLEN & SCHUBÖ (1973) verweisen.

Neuartig hingegen ist die Konfigurationsfrequenzanalyse nach KRAUTH & LIENERT (1973), programmiert von ROEDER (1974, 1976). Da wir uns aus Gründen, die wir sofort darstellen werden, hauptsächlich auf die Konfigurationsfrequenzanalyse (= KFA) stützen, wollen wir sie kurz beschreiben.

Die KFA ist eine Clusteranalyse nach dem 'density-in-space'-Modell (vgl. BAUMANN, 1971). Sie baut also nicht auf dem Assoziationsmaß von je zwei Variablen (= 'inter-id-relations'-Modell; z. B. Faktorenanalyse und herkömmliche Clusteranalyse) auf, sondern auf der Annahme eines bei n Dimensionen n -dimensionalen Raumes. Jede der n Variablen wird nun in $m(n)$ diskreten Ausprägungen gedacht, so daß ein n -dimensionaler Raum mit $n \times m(n)$ Zellen entsteht. Für jede dieser $n \times m(n)$ Zellen wird nun

eine theoretisch zu erwartende Besetzung (E) berechnet, wobei die Randwahrscheinlichkeiten als Schätzung der wahren Parameter zugrunde gelegt werden. Die theoretisch zu erwartende Besetzung (E) wird nun mit der empirisch vorgefundenen (F) verglichen; für diesen Vergleich wird ein Kennwert B berechnet, der nach der üblichen 5-%-, 1-%- usw. Konvention auf eine theoretische Wahrscheinlichkeit bezogen wird. Dadurch ist es möglich, Zellen zu identifizieren, die signifikant unter- bzw. überbesetzt (beim Programm KFA2: überbesetzt) sind. Voraussetzung hierfür ist die Annahme der Unabhängigkeit der verrechneten Variablen, d. h. die Zufallshypothese.

Auf dem Niveau des 'know how' ist die zugrunde liegende Logik nicht anders als beim Vier-Felder-Chi-Quadratstest (vgl. CLAUSS & EBNER, 1968). Dort nimmt man die Unabhängigkeit von zwei dichotomen Variablen an, berechnet über die Schätzung der theoretischen Wahrscheinlichkeiten mit den Randwahrscheinlichkeiten die theoretisch zu erwartende Häufigkeit (E), berechnet für den Vergleich der theoretischen (E) mit den empirischen (F) Häufigkeiten einen (Chi-Quadrat-) Kennwert und vergleicht diesen - unter Berücksichtigung z. B. der 5-%-Konvention - mit einer theoretischen Verteilung (der Chi-Quadrat-Verteilung). In beiden Fällen wird ein möglicher Zusammenhang immer gegen eine Zufalls- (Unabhängigkeits-) Hypothese getestet.

Während beim herkömmlichen Chi-Quadratstest die gesamte $n \times m(n)$ -Tafel gegen den Zufall getestet wird, wird bei der KFA die Differenz zwischen empirischer und theoretischer Häufigkeit jeder einzelnen Zelle gegen die Annahme der Unabhängigkeit geprüft.

Die KFA kann sowohl als Dependenz - wie als Interdependenzanalyse (vgl. SCHULTE, 1972, S. 62ff.) eingesetzt werden: Man kann gleichzeitig den multiva-

riaten Zusammenhang zwischen mehreren Variablen (Interdependenzanalyse) in Abhängigkeit von vorgegebenen Bedingungen (Dependenzanalyse) untersuchen. Im Unterschied zu den parametrischen Verfahren der Interdependenzanalyse (z. B. Faktorenanalyse), der Dependenzanalyse (z. B. Varianzanalyse) und von Mischungen (z. B. Diskriminanzanalyse) sind bei der KFA keinerlei Voraussetzungen für die Datenqualität zu erfüllen.

Ein von KRAUTH & LIENERT (1973) genannter Nachteil der KFA - die im Verhältnis zur Zahl der Merkmale notwendige große Versuchspersonen-Stichprobe - entfällt wegen der von ROEDER (1974, 1976) programmierten simultanen Binomialtests.

Dieser potentielle Vorteil schlägt um in einen Nachteil, weil dadurch die Möglichkeit der Exklusion ungeeigneter, typologisch irrelevanter Variablen entfällt. Wir müssen bereits vor Beginn der Verrechnung wissen, daß die Variablen typologisch relevant sind.

Eine der wichtigsten Beschränkungen für den KFA-Anwender ist weiter der computertechnische Zwang zur Begrenzung der Zahl der simultan verrechneten Variablen. Mit den vorliegenden Programmen können maximal zehn Variablen simultan verarbeitet werden.

Wir mußten deshalb neue Analysewege finden. Sie werden im Zusammenhang der Ergebnisse jeweils einzeln beschrieben werden, unterliegen jedoch alle folgendem Grundprinzip: Zunächst wird jeder Datensatz für

sich verrechnet und beurteilt; die Tests werden z. B. auf faktorielle Validität untersucht, aus der SS werden Leitvariablen extrahiert usw.; zum Schluß werden für jeden einzelnen Datensatz signifikante Typen errechnet. Das Problem der Variablenbeschränkung wird dadurch aufgeschoben, wir erhalten jedoch für jeden Datensatz die typologisch maximal relevanten Items.

Ein Typ in der KFA (vgl. KRAUTH & LIENERT, 1973, S. 30ff.) wird nun dadurch definiert, daß er häufiger (ein Antityp: seltener) auftritt, als aufgrund der Häufigkeit der Einzelmerkmale unter der Nullhypothese ihrer totalen Unabhängigkeit zu erwarten ist. Ein Typ im Sinne der KFA ist also keine besonders auffällige Merkmalsballung (assoziiere: der 'typische' Deutsche) und auch keine besonders häufige Merkmalszusammenstellung (assoziiere: das 'typische' Kompaktauto), sondern eine Merkmalskombination -, die es bei Annahme der Unabhängigkeit der Variablen in dieser Häufigkeit - Konfigurationsfrequenz - nicht geben dürfte. Auch eine absolut sehr niedrige Konfigurationsfrequenz kann also signifikant sein, wenngleich ihre Prägnanz - ihre praktische Bedeutung - dann niedrig sein mag. Die Beurteilung der Überzufälligkeit und der praktischen Bedeutung sind in der KFA zwei verschiedene Schritte.

Das bislang Dargestellte ist vermutlich schneller zu begreifen, wenn wir die KFA auf die herkömmlichen parametrischen Verfahren beziehen. Wir beginnen mit der Darstellung einer üblichen Dependenzanalyse: Wir wollen (fiktiv) annehmen, wir hätten ein zweifaktorielles Design, bei dem ein Treatment vorliegt (1) oder auch nicht (0); die Zellen des Designs könn-

ten wir somit mit 00, 01, 10 und 11 bezeichnen; wir wollen ferner annehmen, daß die abhängige Variable die unterdurchschnittliche Ausprägung a und die überdurchschnittliche Ausprägung A annehmen kann; die Hypothesen sollen lauten, daß beide Treatmentfaktoren die Ausprägung der abhängigen Variablen erhöhen. Für die Varianzanalyse bedeutet dies, daß signifikante Haupteffekte erwartet werden, und dies prüft man, indem man die durch die Faktoren aufgeklärte Varianz in Beziehung zur 'Fehlervarianz' (und nicht zur Gesamtvarianz) setzt. Für die KFA heißt dies, daß man die Treatmentfaktoren und die empirische Ausprägung der abhängigen Variablen als Merkmale der Versuchspersonen verrechnet; man testet also acht Konfigurationen gegen den Zufall: 00a, 00A, 01a, 01A, 10a, 10A, 11a, 11A; obige Hypothese in der Terminologie der KFA lautet: die Konfigurationen 00a und 11A treten überzufällig häufig auf, die Konfigurationen 00A und 11a überzufällig selten. Dabei interessiert nur der Test gegen die Nullhypothese und nicht die Prägnanz der Typen. Letzteres würde - als Analogie zur Varianzanalyse - bedeuten, daß die 'aufgeklärte Varianz' zur 'Gesamtvarianz' in Beziehung gesetzt wird.

Für die Interdependenzanalyse wollen wir den einfachsten Fall der Korrelation annehmen: Wir nehmen an, wir hätten zwei Variablen B und C, die ideal positiv miteinander korrelieren. Wenn wir diese Variablen dichotomisieren, erhalten wir die Ausprägungen b und c (unterdurchschnittlich) und B und C (überdurchschnittlich). In der KFA werden die Konfigurationen bc, bC, Bc und BC auf Signifikanz überprüft. Eine ideale positive Korrelation haben wir dann vor uns, wenn die Konfigurationen BC und bc überzufällig häufig und die Merkmalskombinationen bC und Bc überzufällig selten auftreten. Im weiteren wollen wir nun annehmen, wir wüßten nichts von dieser idealen Korrelation, hätten die Variablen B und C aber zusammen mit der Variablen A erhoben.

Damit hätten wir ein Mischmodell vor uns, in dem einerseits Hypothesen geprüft, andererseits Zusammenhänge a posteriori konstatiert werden müssen. Unter den parametrischen Verfahren finden wir hier kaum eine überzeugende und saubere Lösung, wohl aber bei der KFA. Die KFA kann sowohl für Dependenzanalysen als auch für Interdependenzanalysen eingesetzt werden.

Unser Ziel ist es, zunächst für jeden Datensatz getrennt, Typen zu finden, d. h. Kollektive von Merkmalsträgern, die wir mit möglichst wenigen Variablen beschreiben wollen. Dabei stellt sich das Problem der Auswahl geeigneter Variablen; z. B. haben wir in der Strategiensammlung mehr Variablen als wir in der KFA verrechnen können. Die Vorgehensweise werden wir im Detail bei der Darstellung der Ergebnisse wiedergeben.

Die in jedem Datensatz gefundenen Typen werden sodann 'agglutiniert'. Dazu verrechnen wir zunächst die in je zwei Datensätzen gefundenen Typen in einem gemeinsamen Lauf. Dabei sind zwei wichtige Aspekte zu beachten:

1. Bei der Agglutination zweier Typen kann über den Vergleich der empirischen Häufigkeiten, der Binomialkoeffizienten und der Pränanzkoeffizienten bestimmt werden, ob die Typen dasselbe Kollektiv von Versuchspersonen darstellen oder nicht, ob dasselbe Kollektiv das eine Mal mit dem einen Datensatz erfaßt, das andere Mal mit dem anderen identifiziert wurde (oder nicht).

BREITKOPF (1977b) simulierte in Plasmomen mögliche Ergebnisse. Auf diese Arbeit stützen wir uns, wenn wir die folgenden heuristischen Regeln annehmen:

- 1a) Wir nehmen an, daß es sich um dasselbe Kollektiv handelt, wenn in der gemeinsamen Analyse der Pränanzkoeffizient Q gegenüber den Pränanzkoeffizienten der Einzelanalyse um maximal .05 gesunken, gleich geblieben oder gestiegen ist.

- 1b) Wir nehmen keinen oder einen zufälligen Zusammenhang an, wenn der Prägnanzkoeffizient Q für einen Typ der gemeinsamen Analyse gegenüber den Prägnanzkoeffizienten der Einzelanalysen um mehr als 50 % gesunken ist.
2. Es ist möglich, daß in der gemeinsamen Analyse Kollektive gefunden werden, die in den Einzelanalysen noch als zufällig betrachtet werden mußten. Dies verweist auf einen Vorgehensfehler, wie er bei ROEDER (1976) diskutiert wird. Danach kann eine Folge der in der Praxis der KFA üblichen Dichotomisierung der Variablen sein, daß ein Typ wegen ungeschickter Klassenzuweisung die Signifikanzgrenze 'verpaßt'. Bei der gemeinsamen Analyse zusammen mit neuen Randwahrscheinlichkeiten kann er dann als überzufällig gefunden werden.

Auf diese Weise gefundene Kollektive wollen wir nur dann als relevant betrachten, wenn sie sich in mehreren gemeinsamen Analysen durchsetzen.

Bei diesem Vorgehen tritt jedoch das Problem auf, daß wir einen Datensatz mehrfach verrechnen und damit das nominelle Alpha-Risiko erhöhen. Dieses Problem taucht in jeder KFA auf und wird in der Regel durch eine Alpha-Adjustierung gelöst. Bei der Alpha-Adjustierung wird der Grenzwert der theoretischen Verteilung, der zur Beurteilung der 'Signifikanz' herangezogen wird, nach bestimmten Überlegungen korrigiert (Näheres bei KRAUTH & LIENERT, 1973, S. 39f.) Wir haben in dieser Arbeit immer zur konservativen Maßnahme der totalen Alpha-Adjustierung gegriffen.

Aus theoretischen Gründen müssen wir bei der Analyse gerade der Strategien mit adimensionalen Strukturen rechnen, insbesondere mit Voraussetzungsstrukturen. Der Begriff der Voraussetzung meint dabei, daß die Variation oder Ausprägung der einen Variable

eine bestimmte Merkmalsausprägung einer anderen Variable voraussetzt. Z. B. muß ein Kandidat, der plant, wortreich und ausführlich über sein Spezialgebiet zu reden, voraussetzen, daß nach diesem gefragt wird.

Die Analyse solcher Strukturen ist mit der Voraussetzungsstrukturanalyse von KLEITER (1974) möglich. Dieses Verfahren hat jedoch den Nachteil, daß es keine Korrelationen abbildet. Es gilt demnach, eine statistische Methode zu verwenden, die sowohl Voraussetzungsstrukturen wie Korrelationen verarbeitet.

Dieses Verfahren ist die KFA, die jedoch computer-technischen Beschränkungen unterworfen ist.

Eine Alternative bietet eine Abwandlung der Linkage Analyse nach McQUITTY (1957, 1964), die auch bei LIENERT & KEREKJARTO (1969) beschrieben wird (das Verfahren ist dort so ausführlich beschrieben, daß wir auf eine Darstellung hier verzichten). Bei der Linkage Analyse werden nach einem bestimmten Algorithmus aus einer Assoziationsmaßmatrix Cluster extrahiert.

Wir haben bei den meisten Linkage Analysen den Vierfelder-Chi-Quadratwert als Assoziationsmaß zugrunde gelegt, da dieser Koeffizient auf jedweden Verstoß gegen die Nullhypothese reagiert. Die Verrechnung des Kontingenzkoeffizienten ist dabei nicht notwendig, da dieser mit Operationen aus dem Chi-Quadratwert errechnet wird, die für alle Werte einer Matrix die gleichen sind. Aus der Chi-Quadratmatrix der Variablen extrahieren wir Cluster, die wir konfigurationsfrequenzanalytisch beurteilen und weiterverarbeiten.

Aus pragmatischen Gründen verwendeten wir bei der KFA dichotomisierte oder trichotomisierte Daten.

Die Reduktion von parametrischen Daten auf wenige Klassen ist vom Modellhintergrund der KFA zwar nicht zwingend vorgeschrieben, wird jedoch wegen der in der Praxis sowieso üblichen Reduktion auf wenige entscheidungsrelevante Klassen üblicherweise vorgenommen (vgl. ROEDER, 1974, S. 823).

Ein deswegen gegen die KFA zuweilen vorgebrachter Einwand ist der des Informationsverlustes durch die Dichotomisierung.

Wir meinen jedoch, daß dieses Argument häufig vorschnell geäußert wird. Zum einen kann von einem Informationsverlust nur dann die Rede sein, wenn eine hochreliable, kontinuierliche Skala dichotomisiert wird. Eine Reduktion der Menge der potentiell auftretenden reellen Zahlen ist per se noch kein Informationsverlust in der Psychologie. Zum anderen ist zu überlegen, in welchem funktionalen Zusammenhang die Information verloren geht. Die Dichotomisierung einer hochreliablen Skala, mit der u. U. eine dichotome Entscheidung getroffen wird, sehen wir nicht per se als Informationsverlust.

Demgegenüber stehen Informationsgewinne durch den Einsatz der KFA. Die KFA verarbeitet routinemäßig Korrelationen, Voraussetzungsstrukturen, Tripelassoziationen und Assoziationen noch höherer Ordnung. Mit der Tripelassoziation z. B. ist es möglich, bei Kenntnis zweier Prädiktoren, die nicht miteinander und jeder für sich auch nicht mit dem Kriterium korrelieren, das Kriterium dennoch mit hoher Sicherheit vorherzusagen. Eine Tripelassoziation ist parametrisch nicht erfaßbar (vgl. KRAUTH & LIENERT, 1973, S. 18ff.); sie liegt z. B. dann vor, wenn die Wirkung einer dichotomen Moderatorvariablen in der Umkehrung des Vorzeichens eines Korrelationskoeffizienten besteht.¹ Ohne Berücksichtigung der Moderatorvariablen würden wir eine Nullkorrelation errechnen; mit Berücksichtigung der Moderatorvariablen errechnen wir für das Kollektiv der Merkmalsträger der einen Stufe eine

¹ Einen Hinweis auf diesen Aspekt verdanke ich Herrn Professor Dr. D. Schulte.

negative, das der anderen Stufe eine positive Korrelation. Während bei parametrischen Verfahren die Moderatorvariable vorher bekannt sein muß, wird sie in der KFA rountinemäßig erfaßt.

Ein weiterer Vorteil zeigt sich bei der Behandlung von Missing Values, d. h. der fehlenden Reaktionen. Sie treten z. B. auf, wenn eine Versuchsperson oder ein Proband ein Item eines Fragebogens nicht bearbeitet. Je nach Testkonstruktion wird dies entweder berücksichtigt oder nicht. Manche Tests berechnen aus den Missing-Value-Häufigkeiten den Score für eine eigenständige Subskala, bei anderen Tests werden Missing Values ignoriert und nur die bearbeiteten Items ausgewertet, manche Tests werden bei zu vielen Missing Values als unbrauchbar definiert.

Die KFA bietet nun die Möglichkeit, Missing Values und beantwortete Items in einem Lauf zu verrechnen. Zwar muß die Zuordnung eines Falles zu den m Klassen einer Variable eindeutig sein, die Reihenfolge der Klassen ist jedoch beliebig, d. h. sie braucht nicht einer Rangreihe zu entsprechen. Dadurch ist es möglich, Typen routinemäßig zu identifizieren, die bei bestimmten Items nicht antworten, wohl aber bei bestimmten anderen. Solche Reaktionsmuster sind mit herkömmlichen Verfahren nicht bzw. nur sehr schwer zu isolieren.

Wir können mit der KFA also Ergebnisse erzielen, die mit parametrischen oder herkömmlichen nonparametrischen Verfahren nicht - jedenfalls nicht routinemäßig - zu berechnen sind. Ergebnisse dieser Art stellen einen potentiellen Informationsgewinn dar.

Offen muß beim derzeitigen Forschungsstand freilich bleiben, in welchem Verhältnis der potentielle Informationsgewinn zu einem etwaigen Informationsverlust steht. Hier muß die zukünftige empirische Forschung zeigen, inwieweit der theoretische Vorteil der KFA einen möglichen Nachteil ausgleicht. Im Rahmen dieser Arbeit gehen wir heuristisch davon aus, daß die (eventuelle) Erfassung adimensionaler Strukturen höher zu gewichten ist als der durch die Dichotomisierung entstehende Informationsverlust.

3.5.4 Prinzipielle Planung der Datenanalyse

Durch die Lebensnähe der Erhebung, die eine maximale Standardisierung behindert, einerseits und durch die technische Beschränkung des Hauptcomputerprogramms der KFA andererseits auf maximal zehn simultan verrechnete Variablen sind bestimmte Veränderungen bei der Datenanalyse bedingt.

Zunächst ist der Nachweis zu führen, daß die hypothesenbezogenen Kerndaten keine Fehlervarianz beinhalten, die größer ist als die unter Laboratoriumsbedingungen erhaltene. Ferner ist aufzuzeigen, daß die von uns untersuchte Stichprobe sich nicht wesentlich von den üblichen 'Zufallsstichproben' aus Psychologiestudenten unterscheidet.

Die Peripherdaten lassen die interne Validität (CAMPBELL & STANLEY, 1963, deutsch: 1970) der Kerndaten abschätzen. Die Peripherdaten ersetzen in dieser Arbeit also die uns nicht möglichen Laboratoriumsmaßnahmen zur Erhöhung der internen Validität.

Wenn es uns gelingt, die in den einzelnen Abschnitten wiederholten Hypothesen anzunehmen und bekannte Ergebnisse erneut zu replizieren, dann können wir darauf verweisen, daß die Kerndaten unter genau den gleichen Bedingungen erhoben wurden und die Kerndatenergebnisse deshalb nicht auf Zufallsbedingungen zurückgeführt werden können.

Die 'Darstellung der Ergebnisse' eröffnen wir deshalb mit der Deskription der Stichprobe und der natürlichen Designs sowie mit der Darstellung der Ergebnisse, die sich in den Analysen der peripheren Datensätze ergeben.

Im Anschluß daran stellen wir die hypothesenbezogenen Ergebnisse dar, d. h. diejenigen Ergebnisse, die sich aus den Datensätzen zur Erfassung der Strategien, der Belastungen, der Leistungen und der Vorerfahrung ergeben. Im einzelnen handelt es sich um folgende Datensätze:

1. Strategien und Taktiken,

- die Testverhaltensdifferentiale TVD(W) und TVD(H),
- die Strategiensammlung SS,
- die durch die Beisitzer beobachteten und skalierten Prüfungssteuerungen;

2. Belastungen,

- die Emotionstests STAI, BDI und HiS,
- die Analogskalen,
- die Beschwerdeliste BL.
- die fremdskalierte Aufregung zu Beginn und gegen Ende der Prüfung;

3. Leistungsdaten,

- die erhaltene Zensur,
- der Prozentsatz des geprüften Stoffes, der vor Beginn der Prüfungsvorbereitungen bekannt war,
- die erwartete Zensur,
- die Zensurbeurteilung durch den Beisitzer;

4. Die Klassifikation der Prüflinge entsprechend der Variable Prüfungsart (VD vs. HD).

In Tabelle 6 haben wir noch einmal zusammengefaßt, wann wir welche hypothesenbezogenen Daten erheben. Dabei bedeutet eine retrospektive bzw. prospektive Skalierung, daß sich die Skalierenden an eine vergangene Situation erinnern bzw. sich eine zukünftige Situation laut Instruktion vorstellen sollen;

Tabelle 6: Übersicht über die Erhebung der hypothesenbezogenen Datensätze

Datenquelle Skalierung Fragebogen/ Zeitpunkt	Selbstbeobachtung									Fremdbeobachtung
	retrospektiv			on-line			prospektiv			retrospektiv
	BL	AN	F2/F3	BL	AN	F2/F3	BL	AN	F2/F3	
HIS/STAI/BDI Analogskalen Beschwerdeliste Aufregung		x		x						a)Beisitzer, b)Zuschauer
TVD(J)/TVD(H) SS Prüfungssteuerung	x		x		x		x		x	a)Beisitzer, b)Zuschauer
Prozentsatz d. St. Erwartete Zensur Erhaltene Zensur Validität d. Zensur		x		x	x				x	d-Beisitzer

'on-line' heißt, daß von den Skalierenden eine Reaktion in bezug auf die Jetzt-Zeit verlangt wird. Die Kürzel zu den Fragebögen/Zeitpunkten bedeuten: BL = wird bei der Baseline erhoben, AN = wird im Fragebogen AN erhoben (während der Kandidat auf die Zensur wartet), F2/F3 = wird nach Bekanntgabe der Zensur in dem Fragebogen F2 oder F3 erhoben.

Das letzte Kapitel des Ergebnisteils beinhaltet die Ergebnisse aus denjenigen Datensätzen, die wir ausschließlich zur Hypothesengenerierung einsetzen wollen.

4 DARSTELLUNG DER ERGEBNISSE

4.1 Deskription der Stichprobe und des natürlichen Designs

Abweichend vom Üblichen wollen wir die Zusammensetzung der Stichprobe hier beschreiben. Dies ist legitim, weil

- wir eine Vollerhebung durchführen, so daß das Prinzip der Zufälligkeit der Versuchspersonen-Auswahl verletzt wurde,
- das "Design" letztlich nicht vom Autor bestimmt wurde, so daß "Fehler im Design" nicht auf mangelnde Versuchsplanung zurückgeführt werden können.

4.1.1 Die Mortalität des potentiellen und die Rekrutierung der untersuchten Stichprobe

Wie schon erwähnt (vgl. S. 212), bestand die potentielle Stichprobe bei Drucklegung der Baseline-Fragebögen aus 24 Hauptdiplomanden (= HDn) und 44 Vordiplomanden (= VDn). 5 HDn und 3 VDn traten noch vor Beginn der Prüfungen zurück. Bei 3 VDn ergab sich, daß sie einzelne Wiederholungsprüfungen ablegten; sie wurden wegen mangelnder Meßwiederholbarkeit ausgeschieden. 2 VDn und 4 HDn verweigerten die aktive Teilnahme, nicht jedoch die passive, d. h. sie stimmten einer Beurteilung durch die Beisitzer zu. 2 HDn und 3 VDn gaben an, den Baseline-Fragebogen nicht erhalten zu haben; da ihre Fragebögen von der Post nicht als unzu-

stellbar zurückkamen, wurden sie wie Verweigerer behandelt. 2 VDN und 1 HD verweigerten die Teilnahme während der Datenerhebung ohne Begründung. Ein ausländischer VD, der nach Auffassung des Autors und zufällig anwesender Studenten sich bei der ersten Prüfung als der deutschen Sprache kaum mächtig erwies, wurde ausgeschieden. Ebenso wurde mit einem Studenten verfahren, der vor der ersten Prüfung angab, er sei in psychiatrischer Behandlung und erhalte zur Zeit hohe Dosen Lithium; da Prüfer und Beisitzer einen "ungewöhnlichen" Prüfungsverlauf, Bekannte des Studenten die psychiatrische Medikation auf Befragen bestätigten, wurde auf eine weitere Befragung verzichtet.

Tabelle 7: Die Mortalität der potentiellen Stichprobe bzw. die Selektion der Untersuchungsstichprobe

potentielle Stichprobe	N (%)	
	Ja	Nein
Schwäche	1	3
	-	3
	2	2
	2	3
	1	2
	-	1
Untersuchungsstichprobe	12	20

Mit Chi-Quadrat-Tests überprüften wir die Verteilung der Zurückgetretenen und der Verweigerer auf die Prüfungsarten. Beide Tests fielen nicht signifikant aus. Wir fanden demnach kein Indiz dafür, daß HDn häufiger verweigern oder noch kurzfristig zurücktreten.

Tabelle 8 zeigt die N = 41 Teilnehmer aufgeschlüsselt nach Datenerhebungszeitpunkt, Prüfungsart (VD/HD) und Prüfern (A, B, C, D, E). Dabei bedeutet z. B. der Datenerhebungszeitpunkt P2, daß diese Kandidaten innerhalb ihrer individuellen Sequenz von Prüfungen zum zweiten Mal an unserer Datenerhebung teilnahmen.

Tabelle 8: Verteilung der Kandidaten auf Datenerhebungszeitpunkte und Prüfer

Prüfungsart	Datenerhebungszeitpkt.	Prüfer					Σn	Σn
		A	B	C	D	E		
Vordiplom (VD)	P1	0	12	10	0	7	29	4
	P2	0	4	12	0	10	26	6
	P3	0	0	1	0	3	4	12
	P4	0	0	0	0	1	1	13
	$\Sigma n(VD)$	30	16	23	0	26	95	35
Hauptdiplom (HD)	P1	0	0	2	1	5	8	12
	P2	0	0	0	3	4	7	19
	P3	0	0	0	0	9	9	28
	P4	0	0	0	0	0	0	28
	$\Sigma n(HD)$	12	0	2	4	22	48	41
	Σn	42	16	25	4	48	143	76

Es ergaben sich maximal 4 Datenerhebungszeitpunkte (P1 - P4), d. h. kein Prüfling wählte alle fünf Prüf-

fer. Zu den Zeitpunkten P1 und P2 stand uns noch die gesamte Stichprobe mit $N = 41$ Kandidaten, $n(VD) = 29$, $n(HD) = 12$, zur Verfügung. Zum Zeitpunkt P3 (identisch mit dreifach wiederholter Messung) reduzierte sich die Stichprobe auf $n(VD) = 27$. Zum Zeitpunkt P4 war die Stichprobe auf $N = 19$ ($n(VD) = 10$, $n(HD) = 9$) geschrumpft; i. e. eine vierfach wiederholte Messung liegt für $N = 19$ vor. Der Prüfer D wurde von keinem der teilnehmenden VDn, der Prüfer B von keinem der teilnehmenden HDn gewählt.

An dieser Stelle muß betont werden, daß diese Prüfer trotzdem Prüfungen im Vordiplom bzw. Hauptdiplom abnahmen, nur eben nicht bei Kandidaten, die sich zur Teilnahme an unserer Untersuchung bereit erklärt hatten. Da auch die 'Verweigerer' an der passiven Befragung teilnahmen, konnte kein Prüfer zum Zeitpunkt der Prüfung wissen, ob 'seine Prüfung' anschließend beurteilt würde.

In den Fällen, wo die Spaltenzwischensummen der Tabelle 8 größer sind als die Zeilensummen, sind Doppelprüfungen erfaßt, d. h. es gab Kandidaten, die sich an verschiedenen Tagen in zwei Fächern bei einem Prüfer prüfen ließen.

4.1.2 Exkurs: Ausführliche und exemplarische Darstellung eines KFA-Ergebnisses

Mit dem Programm KFA2 überprüften wir Tabelle 8. Gegen die Nullhypothese der totalen Unabhängigkeit testeten wir die Konfigurationsfrequenzen der Variablen

- Prüfungsart (VD/HD) mit zwei Merkmalsausprägungen,
- Datenerhebungszeitpunkt (P1/P2/P3/P4) mit vier Merkmalsausprägungen und

- Prüfer (A/B/C/D/E) mit 5 Merkmalsausprägungen.

Wir stellen die Ergebnisse (Tabelle 9) ausführlich dar, um dem Leser Gelegenheit zu geben, sich mit der Ergebnisdarstellung von Konfigurationsfrequenzanalysen (KFA) vertraut zu machen.

Tabelle 9: Überzufällig häufige Konfigurationen der Variablen Prüfungsart, Datenerhebungszeitpunkt und Prüfer

signifikante Konfiguration	F	E	B	s.	G	s.
VO/P1/Prüfer B	12	11,4	.030-01	xx	.0545	xxx
VO/P3/Prüfer A	18	7,94	.100-02	x	.0762	xxx
HO/P2/Prüfer D	8	0,85	.300-05	xxx	.0514	xxx
HO/P4/Prüfer A	9	1,93	.110-03	xx	.0519	xxx

F = beobachtete Häufigkeit

E = theoretisch erwartete Häufigkeit

B = Signifikanzzahl für diese Konfiguration

Q = Prägnanzkoeffizient nach KRAUTH & LIENERT (1973)

x = auf dem 5-%-Niveau nach totaler Alpha-adjustierung signifikant

xx = auf dem 1-%-Niveau nach totaler Alpha-adjustierung signifikant

xxx = auf dem 0,1-%-Niveau nach totaler Alpha-adjustierung signifikant

In dem KFA-Lauf wurden 40 Konfigurationen (2 x 4 x 5) gegen die Nullhypothese getestet. Vier davon treten überzufällig häufig auf. Es handelt sich um Kollektive von Merkmalsträgern (Typen), die durch bestimmte Ausprägungen der verrechneten Variablen beschrieben werden können.

So haben wir z. B. ein Kollektiv von Merkmalsträgern mit der gemeinsamen Konfiguration VD/P1/Prüfer B gefunden. Für dieses Kollektiv muß die Nullhypothese der totalen Unabhängigkeit abgelehnt werden, und die Alternativhypothese (es gibt einen Zusammenhang) muß angenommen werden. Die Nullhypothese muß ebenfalls für die anderen signifikanten Konfigurationen (Kollektive) abgelehnt werden.

Inhaltlich heißt dies z. B., daß der Prüfer B zum Datenerhebungszeitpunkt P1 bei den gegebenen Randverteilungen überzufällig viele VDn geprüft hat. Dies heißt nicht, daß er zu viele VDn geprüft hat (vgl. Prüfer C in Tabelle 8), sondern dies heißt, daß der für die Zelle VD/P1/Prüfer B mit Hilfe der Randverteilungen geschätzte theoretische Erwartungswert (E) signifikant kleiner ist als die beobachtete Konfigurationsfrequenz (F).

Die Ergebnisse der KFA lauten zusammengefaßt in verbaler Formulierung: Die Nullhypothese der Unabhängigkeit der Variablen Datenerhebungszeitpunkt, Prüfer und Prüfungsart wird zurückgewiesen, die Alternativhypothese, nach der bestimmte Prüfer bestimmte Gruppen zu bestimmten Datenerhebungszeitpunkten überzufällig häufig prüften, wird angenommen.

Manchen Leser mag die Darstellung der Signifikanzzahl irritieren. Es handelt sich um die aus der Datenverarbeitung übernommene doppeltgenaue Zahlendarstellung (vgl. MÜLLER & STREKER, 1970).

Tabelle 10: Überführung der herkömmlichen in die doppeltgenaue Zahlendarstellung

<u>herkömmliche Zahlendarstellung</u>		<u>doppeltgenaue Zahlendarstellung</u>
0.000 798	≐	.798D-03
0.00798	≐	.798D-02
0.0798	≐	.798D-01
0.798	≐	.798D+00
7.98	≐	.798D+01
79.8	≐	.798D+02
798.0	≐	.798D+03

In Tabelle 10 haben wir erfundene Beispiele dargestellt. Die Zahl hinter dem D gibt an, um wieviel Stellen das Komma nach links oder rechts gerückt werden muß.

Wir wollen zugleich das Prinzip der Alpha-Adjustierung am Typ VD/P3/Prüfer A verdeutlichen. Bei fünf (Prüfern) $\times$ 4 (Datenerhebungszeitpunkten) $\times$ 2 (Prüfungsarten) = 40 Konfigurationsmöglichkeiten ist ein Typ auf dem Signifikanzniveau von $p = .05$ ("fünf Prozent") erst dann signifikant, wenn seine Signifikanzzahl B kleiner ist als $B = .05/40 = .00125 = .125D-02!$

4.1.3 Deskription der untersuchten Stichprobe

Für die untersuchte Stichprobe wurden die demographischen Daten der Tabelle 11 ermittelt (N = 41).

Tabelle 11: Mittelwerte (M), Standardabweichungen (s), t-tests und F-Tests für die demographischen Daten

Variante	M(MD)	M(MD)	c(t-test)	s(VD)	s(MD)	c(F-test)
Alter	23.4	27.9	xxx	2.8	3.6	
Fachsemester	5.0	12.9	xxx	0.9	1.7	xx
Hochschulsemester	6.7	13.3	xxx	3.0	1.9	
Semester						
.. verolencert	1.8	1.5		1.7	1.2	
.. beurlaubt	0.3	0.2		0.8	0.9	
.. effektiven Studiums	3.0	9.3	xxx	1.6	3.3	xx
Zahl der Tage von Ausfüllen des Fragebogens bis zur 1. Prüfung	15.6	19.7		6.3	6.5	
Datum, an dem der Fragebogen ausgefüllt wurde (Tag)	20.7	17.9		3.9	6.8	xx
Zahl der bis jetzt an einer Universität abgelegten Prüfungen	1.6	6.3	xxx	4.9	1.1	xx
Zahl der Prüfungen im Abitur	1.5	1.4		2.0	1.0	x
Zahl der Prüfungen, die ...						
.. im Arbeitsaufwand vergleichbar waren	1.2	5.3	xxx	3.2	2.5	
.. in der Schwere vergleichbar waren	0.5	4.7	xxx	1.6	2.7	x
Dem Prüfer aufgefallen durch ...						
.. Seminare Diskussionen	0.3	0.4		0.5	0.5	
.. schriftliche Arbeiten	0.1	0.4	xx	0.3	0.5	xx
.. politische Ansichten	0.2	0.2		0.4	0.4	

Alle statistischen Tests, bei denen signifikante Unterschiede zwischen VDn und HDn zu erwarten waren, fielen entsprechend aus. HDn sind signifikant älter, haben sowohl mehr Fach- als auch Hochschulsemester absolviert, haben mehr Semester effektiv studiert und haben mehr Prüfungen an einer Universität bis zu diesem Zeitpunkt abgelegt, auch solche, die in bezug auf Arbeitsaufwand oder Schwere mit den anstehenden vergleichbar waren. HDn glauben signifikant häufiger (χ^2 -Test), sie seien durch schriftliche Arbeiten aufgefallen.

Wenig überraschend ist auch, daß die Zahl der Tage bis zur ersten mündlichen Prüfung sowohl in der expliziten Frage wie im indirekten Zugang über das Bearbeitungsdatum keinen Gruppenunterschied ergibt. Die Zahl der mündlichen Prüfungen im Abitur ist ebenfalls annähernd gleich. Interessanter ist der fehlende Unterschied in den Angaben zu verplemperten und beurlaubten Semestern: die durchschnittlichen Angaben beider Gruppen gleichen sich.

Tabelle 12: Verteilung der Geschlechter in der Stichprobe

Verteilung der Geschlechter			
	VD	HD	Σ
männlich	15	10	25
weiblich	14	2	16
Σ	29	12	41

Die Verteilung der Geschlechter in beiden Gruppen fiel überraschend aus: nur **zwei Frauen legten die Hauptdiplomprüfungen** ab. Deshalb bestimmten wir einen möglichen Einfluß des Geschlechts mit einer eher groben Methode; wir wiederholten den oben dargestellten Gruppenvergleich hinsichtlich der demographischen Variable einmal nur für die Männer, zum anderen berechneten wir innerhalb der Vordiplomanden die Geschlechtsunterschiede jeweils mit t-tests. Dabei ergab sich eine Replikation sowohl der signifikanten als auch der nicht-signifikanten VD/HD-Unterschiede auf dem jeweiligen Niveau mit Ausnahme von 2 Variablen. Die männlichen VDn gaben auf dem 5-%-Niveau signifikant weniger Tage bis zur ersten mündlichen Prüfung an als die männlichen Hauptdiplomanden, während in den Angaben zu den bislang abgelegten mündlichen Prüfungen kein Unterschied mehr berechnet werden konnte. Der Vergleich der Geschlechter innerhalb der VDn-Gruppe ergab keinerlei signifikante Differenz.

Wir gehen im folgenden von der Annahme aus, daß sich ein starker Geschlechtseinfluß stärker in den Ergebnissen niederschlagen müßte als oben dargestellt, so daß die demographischen Daten von Geschlechtseinflüssen frei betrachtet werden können. Wir werden Geschlechtsunterschiede im folgenden daher vernachlässigen.

4.1.4 Der tatsächliche Prüfungsverlauf

Die Kandidaten hatten während des Wartens auf die Zensur den Fragebogen AN auszufüllen. Dies geschah

bei drei Prüfern in deren Sekretariaten, bei zwei Prüfern vor ihren Zimmern. Nach Bekanntgabe der Zensur wurden den Kandidaten die Fragebögen F2 und F3 übergeben, und sie wurden aufgefordert, F2 sofort auszufüllen, da für die dort verwendeten Variablen eine "frische Erinnerung" nötig sei; F3 konnte nach der ursprünglichen Planung mit nach Hause genommen werden.

Wie vom Autor vorhergesehen, traten bei diesem Procedere Schwierigkeiten auf, die eventuell die Fehlervarianz erhöhten, aus organisatorischen und institutionellen Gründen jedoch unvermeidbar waren.

Das schwerste Handicap ergab sich aus der zeitweiligen Parallelität von Prüfungen. Wir betrachteten es als unerlässlich, die Fragebögen persönlich zu übergeben, um die Bearbeitung zu kontrollieren und die Teilnahmemotivation zu erhalten. Für parallele Prüfungen warben wir drei Studenten aus dem Grundstudium an, trainierten jeden drei Stunden und setzten sie ein.¹ Einer der Helfer verhielt sich jedoch nicht instruktionsgemäß! Da es sich als unmöglich erwies, die genaueren Umstände aufzuklären, werden die Daten von 9 Kandidaten dieser Prüfungen teilweise nicht mit verrechnet. An den notwendigen Stellen werden wir darauf verweisen.

Das Angebot, F3 "ungestört" zu bearbeiten, nutzten nur 3 Kandidaten aus. Die überwältigende Mehrheit

¹ Wir danken an dieser Stelle Herrn Detlev Meier, Frau Andrea Plewka und Frau Edith Scherello.

zog es vor, "alles gleich hinter sich zu bringen". Interviews der drei ergaben, daß diese stets unter den gleichen Umständen und stets nach der gleichen Zeitdifferenz die Fragebögen bearbeiteten. Die Daten dieser Kandidaten wurden daher nicht gesondert bearbeitet.

Die für die Bearbeitung von AN kalkulierten 90 Sekunden erwiesen sich als realistisch. In nur 9 von 140 Fällen waren Prüfer und Beisitzer schneller. Nur in einem Fall hatte der Kandidat keine Gelegenheit, AN vor Bekanntgabe der Zensur vollständig zu bearbeiten. Dieser Kandidat bearbeitete die letzten drei Analogskalen nach Bekanntgabe der Zensur.

4.1.5 Exkurs: Unterschiede zwischen Teilnehmern und Verweigerern

Nach McDAVID & HARARI (1968, S. 403 ff.) konnten wir nicht ausschließen, daß sich die Kandidaten, die eine Teilnahme an unserer Erhebung explizit verweigerten, von den Teilnehmern unterschieden. Wir legten die Beisitzer-Daten zugrunde, d. h. die Ergebnisse der "passiven" Befragung, und testeten zweiseitig.

Aus technischen Gründen mußten wir uns auf eine Taschenrechnerauswertung von Zwischenergebnissen beschränken.

Danach fanden sich keine signifikanten Unterschiede zwischen Teilnehmern und Verweigerern in der Beurteilung hinsichtlich des Ausweichens beim Antworten

und des expliziten Zurückweisens von Fragen, der Unterbrechungen durch den Prüfer und einer abweichenden Einschätzung der Prüfungsleistung durch den Beisitzer.

Signifikant unterschiedlich war die Häufigkeit, mit der der Beisitzer beim Prüfer eine Emotion beobachten konnte. Die Vierfeldertafel (Verweigerer/Teilnehmer; nichts beobachtet/mindestens 1 Emotion beobachtet) ließ einen Chi-Quadratwert von $\chi^2 = 5.951$ errechnen, der bei $df=1$ auf dem 5-%-Niveau signifikant ist. Zu diesem Wert lieferten die Daten der Verweigerer den größten Beitrag: die Konfigurationsfrequenz der Zelle Verweigerer/Emotion beobachtet ist (fast) auf dem 5-%-Niveau überbesetzt ($\chi^2 = 3.803$, $df=1$).

Eine Analyse der einzelnen Prüferemotionen entfiel wegen zu geringer Fallzahlen.

t-tests der skalierten Aufregung zu Beginn und zu Ende der Prüfungen ergaben für die einzelnen Datenerhebungszeitpunkte die Ergebnisse der Tabelle 13.

Tabelle 13: Mittelwerte und Signifikanzniveaus der t-tests zwischen Verweigerern und Teilnehmern an der Untersuchung bei den fremdskalierten Variablen "Aufregung zu Beginn" und "Aufregung gegen Ende"; beim Datenerhebungszeitpunkt P2 wurden Missing Values berücksichtigt

Datenerhebungszeitpunkt	N		Aufregung zu Beginn			Aufregung gegen Ende		
	Verweigerer	Teilnehmer	M(Verw.)	M(Teiln.)	p	M(Verw.)	M(Teiln.)	p
P1	11	41	5.13	5.40	<.10	5.45	4.40	<.20
P2	10	32	4.90	4.69	>.40	4.70	4.50	>.40
P3	5	39	4.57	3.95	<.20	3.67	3.27	<.40
P4	5	19	5.40	4.90	<.10	4.20	3.55	<.10

Mit Ausnahme der "Aufregung gegen Ende" zum Zeitpunkt P3 sind alle Mittelwerte der Teilnehmer niedriger als die der Verweigerer, die Aufregung zu Beginn der Zeitpunkte P1 und P4 sogar auf dem 10-%-Niveau.

Bei den Verweigerer-Daten fanden sich keine signifikanten Unterschiede zwischen Vor- und Hauptdiplomanden.

Insgesamt finden sich demnach Indizien, daß Verweigerer beim Prüfer mehr Emotionen auslösen und vermutlich im Schnitt aufgeregter sind als die Kandidaten, die an unserem Versuch teilnahmen.

4.1.6 'flaps'

Mit 'flaps' werden spektakuläre Fälle des 'Ausflip-pens' oder der 'Blockiertheit' kurz vor oder während der Prüfung bezeichnet, die die Assoziationen von Studenten zur 'Prüfungsangst' bestimmen. Während unserer Datenerhebung konnten wir (allerdings nur in einer Prüfung) einen solchen Fall beobachten.

Es handelt sich um eine Studentin durchschnittlichen Alters, die die Vordiplomprüfung ablegen wollte. Sie erschien ca. 20 Minuten vor Beginn ihrer Prüfung und wartete zusammen mit anderen Prüfungskandidaten. Die Gruppe diskutierte, wobei potentielle Prüfungsfragen und die vorhergehenden Prüfungen den Hauptteil der Themen stellten. Ca. 5 Minuten vor Beginn ihrer Prüfung wurde ihr von einem anderen Prüfungskandidaten eine Frage nach einer Definition gestellt, die vermutlich in der anstehenden Prüfung auftauchen würde. Die Studentin fing daraufhin an zu jammern, sie werde die Prüfung wohl nie bestehen, und begann, mit schnell-

len und ruckhaften Bewegungen in ihren Unterlagen vor- und zurückzublättern. Nach ca. 30 Sekunden brach sie in Tränen aus, äußerte, jetzt sei alles vorbei. Ca. 60 Sekunden lang weinte sie nun und äußerte in unvollständigen Sätzen Inhalte, die der Autor im Sinne von 'Ich werde durchfallen' und 'Ich weiß gar nichts mehr' verstand. Nach einer letzten Äußerung 'Oh je, wenn mich ... (Prüfer) jetzt sieht, au weia ...' lief sie zur Damentoilette. Zwei Kommilitoninnen folgten ihr. Diese berichteten, die Studentin habe sich 'im Klo verbarrikadiert', würde weinen und habe geäußert, man möge dem ... (Prüfer) ausrichten, sie sei nicht erschienen. Die Situation blieb unverändert, bis der Prüfer die Studentin zur Prüfung holen wollte. Er erfuhr von dem Geschehen und ließ der Studentin durch seine Sekretärin ausrichten, man müsse dann doch wenigstens einen neuen Termin ausmachen, sie solle zu diesem Zweck in sein Zimmer (das Prüfungszimmer) kommen. Die Studentin folgte - immer noch weinend - dem Prüfer in dessen Zimmer. Ca. 25 Minuten später verließ sie es wieder - weinend - und berichtete, sie habe eine '3' bekommen. Der Beisitzer berichtete später, der Prüfer habe zunächst beruhigend auf die Studentin eingeredet, habe ihr dann einen Termin vorgeschlagen, habe wieder beruhigend auf sie eingeredet und habe sie dann - ohne sie die ganze Zeit zu Worte kommen zu lassen - gefragt, was sie (bei dem zukünftigen Termin) antworten würde, wenn er sie 'beispielsweise' fragen würde ...

4.1.7 Zusammenfassung und Übersicht

Die untersuchte Stichprobe besteht aus 29 VDN und 12 HDN. Es ergab sich die Gelegenheit, bis zu viermal die Messungen zu wiederholen.

Keiner der Verweigerer (in bezug auf die aktive Teilnahme an unserer Untersuchung) weigerte sich, sich durch den Beisitzer beurteilen zu lassen, verweigerte also auch die 'passive Teilnahme'.

Beim Vergleich der demographischen Daten aus der Baseline-Erhebung zwischen VDn und HDn fielen alle Vergleiche in den formalen Variablen wie Semesterzahl, Alter etc. signifikant in der erwarteten Richtung aus. Geschlechtsunterschiede fanden sich nicht; deswegen gehen wir im folgenden davon aus, daß Geschlechtsunterschiede vernachlässigt werden können.

Im Verlauf der Datenerhebung konnten die genauen Umstände bei der Erhebung einzelner Datensätze nicht mehr geklärt werden. Diese Datensätze werden ggf. als Missing Values behandelt.

Wir fanden Indizien dafür, daß Verweigerer im Schnitt aufgeregter sind als die Teilnehmer unserer Untersuchung und daß sie vermutlich beim Prüfer mehr Emotionen auslösen.

Spektakuläre Fälle des 'Ausflippens' konnten wir nur einmal beobachten; negative Folgen für die Kandidatin entstanden dadurch nicht, ebensowenig wie sie das Verhalten wiederholte.

Fälle von Erkrankungen, die so stark waren, daß die Prüfungssequenz abgebrochen werden mußte, traten in dem von uns beobachteten Prüfungszeitraum und in der von uns untersuchten Stichprobe nicht auf.

4.2 Zur internen Validität der erhobenen Datensätze: Ergebnisse der peripheren Datensätze

4.2.1 Der somatische Zustand:

Die Freiburger Beschwerde-Liste (FBL)

Bei der Baseline-Erhebung wurde den Versuchspersonen die Freiburger Beschwerde-Liste (FAHRENBURG, 1976) gegeben. Die Originalfragen wurden durch Kurztexte ersetzt (vgl. S. 174f.) die Originalitems 10, 11, 20, 22, 44, 47, 51 wurden nicht verwendet.

Um die Daten mit FAHRENBURGS Daten vergleichbar zu machen, wurden die Subskalen nach den Angaben des Testautors berechnet und - sofern ein fehlendes Originalitem berücksichtigt werden mußte - die Summenscores mit der Operation (Score/Zahl der verwendeten Items) x Zahl der Originalitems) geschätzt. Das Ergebnis ist in Tabelle 14 dargestellt.

Mit t-tests für unabhängige Stichproben verglichen wir die Prüfungskandidaten mit zwei Eichstichproben von FAHRENBURG (1976).

Von einem gesicherten Unterschied zu den Eichstichproben wollen wir dann sprechen, wenn der Vergleich mit beiden Stichproben signifikant ausfällt. Die Unterschiedshypothese wollen wir dann zurückweisen, wenn keiner der Vergleiche signifikant ausfällt.

Tabelle 14: Mittelwerte und Standardabweichungen der Freiburger Beschwerde-liste (FBL) sowie die Vergleichsdaten der Eichstichprobe

Subskala	Arithm.Mittelwert		Standardabweichung		t-tests zum Vergleich unserer Stichprobe mit 1. St. pr. 2. St. pr.			
	unserer Stichprobe (N=41)	Fahrenbergs 1. Stichpr. (N=330)	Fahrenbergs 2. Stichpr. (N=400)	unserer Stichpr. 1. Stichpr. 2. Stichpr.	Fahrenbergs 1. Stichpr. 2. Stichpr.	unserer Stichprobe mit 1. St. pr. 2. St. pr.		
FBL 1: Allgemeinbefinden	23.14	20.34	18.71	5.98	6.94	5.09	xx	xxx
FBL 2: Emotionale Reaktivität	17.25	19.60	16.26	4.93	6.30	4.53	x	
FBL 3: Herz - Kreislauf	15.97	16.67	12.59	6.56	6.90	4.04		xxx
FBL 4: Magen - Darm	17.34	16.43	14.91	5.82	6.37	4.48		xxx
FBL 5: Kopf - Hals Reizsyndrom	17.83	15.51	16.60	6.50	5.25	4.86	xx	
FBL 6: Anspannung	21.55	17.16	17.38	6.47	6.67	5.20	xxx	xxx
FBL 7: Sensorik	18.28	19.04	18.24	5.92	6.61	5.37		
FBL 8: Schmerz	18.00	15.45	13.00	6.77	6.02	4.30	xx	xxx
FBL 9: Motorik	16.90	14.36	15.53	5.12	5.01	5.23	xx	
FBL 10: Haut	13.29	12.84	11.96	5.62	4.83	3.72		x
FBL 11: Beschwerden-summe	179.34	167.37	155.18	46.08	45.38	33.74		xxx

Überdurchschnittliche Werte erreichen die Prüfungskandidaten demnach in den Subskalen FBL 1 (Allgemeinbefinden), FBL 6 (Anspannung) und FBL 8 (Schmerz).

Beispiele für FBL 1: morgens müde und zerschlagen,
schnelles Ermüden, Benommenheit,
Kopfschmerzen, Appetitmangel

Beispiele für FBL 6: diverse Schlafschwierigkeiten,
feuchte Hände, Schweißausbrüche,
Verkrampfungen

Beispiele für FBL 8: Schmerzen in Nacken, Schultern,
Kreuz und Beinen; in den Extremitäten Kribbeln, Prickeln, Taubheit

Insgesamt fallen Prüfungskandidaten in gesichertem Ausmaß durch eine Symptomatik auf, wie sie SCHULTE & TÖLLE (1973) als "psychovegetatives Erschöpfungssyndrom" beschreiben. Zu beachten ist dabei, daß die Daten 14 Tage vor der ersten mündlichen Prüfung erhoben werden und von daher anzunehmen ist, daß die Prüflinge ihren Lebensstil voll auf die Prüfungsvorbereitungen eingestellt haben.

Unsere Hypothese, Prüfungskandidaten unterschieden sich nicht von der Normalbevölkerung, muß zurückgewiesen werden. Prüfungskandidaten haben 14 Tage vor der ersten mündlichen Prüfung ein gestörtes Allgemeinbefinden, sind angespannter als andere und empfinden häufiger diffuse Schmerzen oder diffuses Kribbeln in der Muskulatur. Verglichen mit anderen

Studenten (2. Eichstichprobe) erreichen Prüflinge höhere Werte in den Herz-Kreislauf-Magen-Darm-Items.

Zurück bleibt eine gewisse Skepsis gegenüber der FBL. Eine Liste, in der sich schon die Eichstichproben signifikant unterscheiden, muß kritisch verwendet werden. Eine von uns durchgeführte Faktorenanalyse konnte die von FAHRENBURG (1976) berichtete Faktorenstruktur auch nicht ansatzweise oder teilweise replizieren.

Auf die Darstellung unserer Lösung verzichten wir, da wir weit weniger Versuchspersonen als Items verrechneten.

4.2.2 Replikation der Struktur des Arbeitsverhaltens: Analyse des Studiengewohnheitsinventars (SGI)

In der Baseline-Erhebung erfaßten wir mit 50 Items das Arbeitsverhalten. Diese Items hatten wir nach MOELLER & SCHEER (1970), TURKSON (1970), SUAREZ (1971) und NAEF (1971) zusammengestellt. Wir gingen davon aus, daß eine Faktorenanalyse dieser Items ähnliche Faktoren ergeben müßte, wie sie in o. g. Arbeiten gefunden wurden. Wegen der im Vergleich zur Itemzahl geringen Stichprobengröße konnten wir von vornherein nicht mit einer stabilen Lösung rechnen; wir verzichteten deshalb von vornherein auf Ähnlichkeitsrotationen o. ä. Andererseits sollten sich mehrfach gefundene Dimensionen wie

- a) Sicherheit und Unsicherheit in der Arbeit,
- b) große bzw. geringe Stetigkeit in der Arbeit und
- c) emotionale Abhängigkeit bzw. Unabhängigkeit

sowie die Assoziation der Dimension a) mit depressiven Merkmalen und Kontaktstörungen und der Dimension b) mit Bequemlichkeit und Unkonzentriertheit (vgl. z. B. MOELLER & SCHEER, 1970) auch in unserer Lösung durchsetzen. Für den Nachweis, daß unsere Daten kein "Rauschen" enthalten, wählten wir die inhaltliche Ähnlichkeit unserer Faktoren mit o. a. Dimensionen als Prüfungskriterium.

Wir berechneten zunächst eine Hauptkomponentenanalyse vom Typ R mit anschließender Varimaxrotation. Entsprechend dem Ergebnis des Scree-Tests entschieden wir uns für eine 6faktorielle Lösung. Alle Items, die in dieser Lösung keine Ladung von $a > |.40|$ auf irgendeinem Faktor erreichten, schieden wir aus und wiederholten die Faktorenanalyse.

Ausgeschieden wurden die Items 5, 15, 19, 21 und 41. Eigenwertverlauf und Faktorenstruktur der neuen Lösung sind in Abbildung 12 und in Tabelle 15 dargestellt.

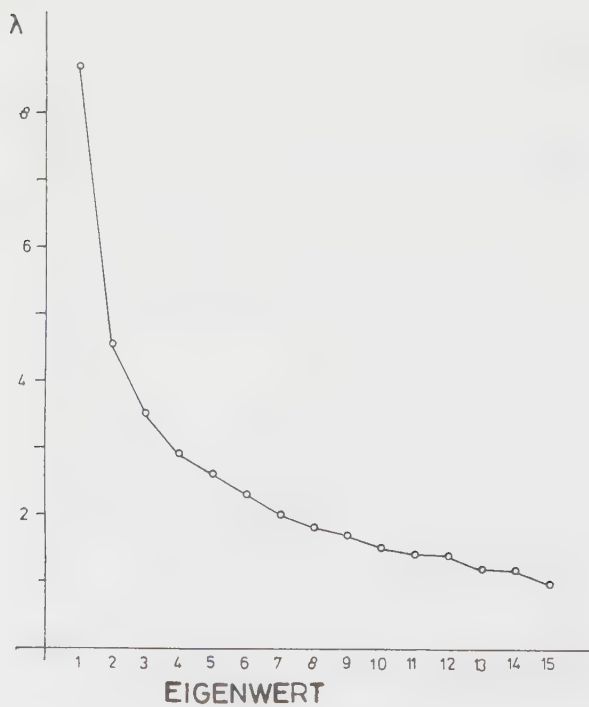


Abbildung 12: Eigenwertverlauf der 2. Faktorenanalyse Studiengewohnheitsinventars (SGI) bei N=41. 15 Eigenwerte > 1.0 bei 6 extrahierten Faktoren

Tabelle 15: 6faktorielle Lösung der varimaxrotierten R-Hauptkomponentenlösung der Items des Studiengewohnheitsinventars bei 54,2 extrahierter Varianz

Faktor 1	Faktor 2	Faktor 3	Faktor 4	Faktor 5	Faktor 6	Text
.81						41. Wenn ich in letzter Zeit merke, daß ich etwas nicht begriffen habe, werde ich sehr unsicher
.71						45. Wenn ich höre, was andere Kandidaten schaffen, werde ich sehr unsicher
.71						47. Ich werde durch die Schwarzmalerei anderer beunruhigt
.67	.43					26. Ich teile mir meine Arbeit häppchenweise ein, dennoch habe ich selten den Eindruck, vorwärts zu kommen
.63						37. Manchmal Überlege ich, ob ich mich auf die Prüfungen richtig vorbereite
.55						31. Besonders in letzter Zeit habe ich den Eindruck, daß die Prüfungspflichtlektüre nur ganz langsam weniger wird
.51						36. Mein Gedächtnis läßt in letzter Zeit zu wünschen übrig
.45					.47	48. Es fällt mir schwerer als anderen, den Stoff zu behalten
-.41				-.59		50. Ich bin der Ansicht, daß 14 Tage Vorbereitung pro Prüfung voll ausreichen
-.51		.46				18. Vor dem Lesen überfliege ich ein Kapitel in einem Lehrbuch
-.65						1. Ich glaube, daß meine persönliche Methode der Vorlesungsverarbeitung wirkungsvoll ist und den Studienanforderungen genügt
	.73					24. Persönliche Probleme oder Tagesereignisse belasten mich in letzter Zeit so stark, daß ich mich kaum auf die Arbeit konzentrieren kann
	.70					7. Es kam häufiger vor, daß ich einen großen Teil einer Vorlesung in Gedanken bei etwas anderem verbrachte
	.48					49. Neben dem, was ich sonst tue, bleibt mir kaum Zeit, mich auf die Prüfungen vorzubereiten
	.57	-.41				4. Es fällt mir schwer, mit der Arbeit anzufangen
	.46					11. Auf meinem Schreibtisch liegen Dinge, die mich zuweilen vom Lernen ablenken
	-.42				.44	6. Wenn ich Lehrbücher oder Zeitschriftenartikel lese, mache ich mir regelmäßige Notizen
	-.49			.54		17. Wenn ich Schwierigkeiten mit dem Stoff habe, frage ich um Rat
	-.51					46. Insgesamt glaube ich, mich intensiver als andere vorzubereiten
	-.52	.60				25. Ich habe meine festen Arbeitszeiten vorher festgelegt und halte mich i.A. an diesen Plan
	-.69					3. Genau wie früher werde ich auch jetzt mit meiner Arbeit meist fertig

Faktor 1	Faktor 2	Faktor 3	Faktor 4	Faktor 5	Faktor 6	Text
		.72				22. Morgens vor Arbeitsbeginn frühstücke ich ausgiebig und gemütlich
		.69				43. Zuweilen macht mir das Lernen Spaß
		.56				10. Ich habe einen festen Arbeitsplatz für meine Prüfungsvorbereitungen
		.51				13. Bei einem Zeitschriftenartikel lese ich in jedem Fall erst die Zusammenfassung, bevor ich den Text durcharbeite
		.42			-.40	32. Ich bin mir völlig sicher, daß ich den Stoff für einige Prüfungen schaffe
		-.44				39. Ich bin unsicher, ob ich auch nur eine Prüfung so weit vorbereiten kann, daß ich wenigstens die Pflichtlektüre gelesen habe
		-.54				20. Beim Lesen lasse ich Tabellen und graphische Darstellungen unbeachtet
			.79			35. Ich bereite mich auf die Prüfungen allein vor
			.40	.55		12. Wenn ich bei einem englischen Text auf ein Wort stoße, das ich nicht kenne, schlage ich es im Wörterbuch nach
			-.62			44. Ich lasse mich bei den Prüfungsvorbereitungen abhören
			-.74			27. Der Stoff, den ich in diesen Tagen durcharbeite, ist mir zumindest in Umrissen meist schon bekannt
			-.83			40. Meine Prüfungsvorbereitungen beinhalten Diskussionen in einer Arbeitsgruppe
				.73		33. Ich mache alle 1 1/2 - 2 Stunden Arbeitspausen
				.43		23. Zur Zeit suche ich häufig die Institutsbibliothek auf
				-.49		38. Mir fällt es nicht schwer, 4 - 5 Stunden ununterbrochen zu lernen
				-.59		29. Pausen zwischen Arbeitsblöcken kann ich mir in diesen Tagen nicht leisten
				-.66		14. Bei den Prüfungsvorbereitungen versuche ich, so viel wie möglich auswendigzulernen
					.69	8. Wenn ich mir in einer Vorlesung Notizen machte, habe ich meist so schnell wie möglich geschrieben
					.64	9. In Vorlesungen, Übungen und Seminaren bin ich zuweilen so sehr mit meinen Notizen beschäftigt, daß ich den Faden verliere
					.55	28. Manchmal merke ich, daß ich etwas wohl visuell gelesen, geistig aber nicht aufgenommen habe
					.49	30. Sehr häufig könnte ich alle Bücher gegen die Wand werfen
					-.50	16. In Seminaren diskutiere ich zuweilen eifrig mit

In dieser Analyse herausgefallen: Items 2, 34

Vorher ausgeschieden: Items 5, 15, 19, 21, 42

In dieser Analyse erreichten die Items 2 und 34 keine Ladung $a > |.40|$ auf irgendeinem Faktor; sie blieben deshalb unberücksichtigt.

Faktor 1: Unsicherheit bzw. Sicherheit in der Arbeit

Auf diesem bipolaren Faktor laden auf dem einen Pol Items, die Unsicherheit und Beunruhigung thematisieren; Items, die gelassene Sicherheit gegenüber Anforderungen beinhalten, auf dem anderen Pol.

Faktor 2: Stetigkeit und Konzentriertheit bzw. Unstetigkeit und Ablenkbarkeit in der Arbeit

Auf diesem bipolaren Faktor laden Items, die unter dem Begriff "Ablenkbarkeit" subsummiert werden können, auf dem einen Pol; Items, die konzentriertes und stetiges Arbeiten erfassen, auf dem anderen Pol.

Faktor 3: Adäquates bzw. inadäquates Arbeitsverhalten (allgemein)

Obwohl auch dieser bipolare Faktor Momente der Unkonzentriertheit enthält, finden wir hier mehrheitlich Items, die adäquates oder inadäquates Arbeiten definieren.

Faktor 4: Soziales Arbeitsverhalten

Mit Ausnahme von Items 17 (... frage ich um Rat) enthält dieser Faktor alle Items, die das Problem Gruppen- versus Einzelarbeit thematisieren.

Faktor 5: Adäquates bzw. inadäquates Arbeitsverhalten (konkret)

Auf diesem bipolaren Faktor laden - im Gegensatz zu Faktor 3 - alle Items, die konkrete Angaben hinsichtlich adäquaten bzw. inadäquaten Arbeitsverhaltens beinhalten.

Faktor 6: Impulsivität und Zerrissenheit

Auf diesem bipolaren Faktor laden Items, die Aspekte der unkontrollierten Impulsivität erfassen, auf dem einen Pol.

Insgesamt haben wir eine plausible Faktorenstruktur gefunden. Die Dimensionen der Unstetigkeit und der Unsicherheit betrachten wir als überzeugend repliziert.

4.2.3 Replikationen des Angstverlaufs

Die Analyse der Emotionsverläufe soll sicherstellen, daß die Fehlervarianz auch während der anderen Datenerhebungszeitpunkte eine Quantité négligeable darstellt.

Wir wissen, daß die Angst unmittelbar vor einem kritischen Ereignis (z. B. einem Fallschirmsprung oder einer Prüfung) ein Maximum erreicht und dann permanent zurückgeht (vgl. S. 178f.) Dieser Mechanismus sollte bei allen Datenerhebungszeitpunkten zu finden sein.

Ferner wissen wir, daß bei Fallschirmspringern ty-

pische Unterschiede zwischen Erfahrenen und Unerfahrenen auftraten: bei Unerfahrenen findet der Forscher eine stärkere Angst unmittelbar vor dem Sprung als bei Erfahrenen, während Erfahrene kurz vor der Landung stärkere Angst empfinden. Im Überblicksartikel von EPSTEIN (1977) wird gezeigt, daß dies sowohl für simulierte Situationen (durch Anbieten von Reizmaterial unterschiedlicher Intensität) als auch für Realsituationen gilt. Für den Problembereich der Prüfungsangst wurde eine Simulation von PIEHL (1972) durchgeführt.

Diesbezügliche Daten erhielten wir bereits bei der BL-Erhebung. Auf die Frage nach Angst im Zusammenhang mit der Mehrheit der mündlichen Prüfungen erhielten wir die Ergebnisse der Tabelle 16.

Tabelle 16: Angaben aus der Baseline-Erhebung zur Prüfungsangst.

Kandidaten	VD (n=29)	ND (n=12)
Angst..		
lange vorher	8	1
unmittelbar		
vorher	15	11
während	6	4

HDn ist das Phänomen der Angst unmittelbar vorher also in 91,7 % aller Fälle bekannt gegenüber nur 51,7 % bei VDn; der Unterschied ist auf dem 5-%-Niveau signifikant.

Einen Zugang zu den Emotionsverläufen bieten uns die Messungen mit den Analogskalen. Wir hatten direkt nach der Prüfung - während des Wartens auf die Zensur - die Emotionen gemessen und nach der Bekanntgabe der Zensur die Emotionen "vor" und "während" der Prüfung retrospektiv skalieren lassen.

Tabelle 17 gibt die Mittelwerte der Analogskalen wieder. Die routinemäßig berechneten Mittelwertsunterschiede zwischen VDn und HDn fielen in nur 2 von 90 Fällen (= 2,22 %) signifikant aus.

Verlaufsderstellung (Mittelwerte)											
An.sc.	Zpkt.	M(VD)	M(HD)	n($\bar{L}_0$)	s.	n($\bar{L}_0$)	s.	n($\bar{L}_0$)	s.	Summe s.	
	p3-vor	6.00	6.58	17	x	14	8			3.73 7.54 x	
	p3-währ.	5.26	6.25	21		8	10				
	p3-nach	4.37	3.92								
	p4-vor	6.36	6.11								
	p4-währ.	6.20	5.67								
	p4-nach	5.27	4.11								
Freude	BL	--nicht erhoben--									
	p1-vor	2.76	3.00							2.92 15.84 xxx	
	p1-währ.	3.79	3.33	10	11	11	18				
	p1-nach	4.62	4.00	5	x	9	24	xx			
	p2-vor	2.45	2.42	12	11	11	16				
	p2-währ.	3.00	3.00	1	xx	7	23	xxx	++		
	p2-nach	3.93	3.73								
	p3-vor	3.12	2.83	10	16	13	13			1.08 25.03 xxx	
	p3-währ.	3.55	3.08	7		12	20				
	p3-nach	4.41	4.33								
	p4-vor	3.27	3.11								
Hilflos.	p4-währ.	2.53	3.89							1.38 6.62 x	
	p4-nach	4.82	3.56								
	BL	4.38	3.17								
	p1-vor	5.17	4.92								
	p1-währ.	4.96	4.75	17	12	10	10			2.00 3.53	
	p1-nach	4.38	4.45	18	11	9					
	p2-vor	5.55	5.58	19	12	8					
	p2-währ.	4.89	5.08	12	9	10			+++		
	p2-nach	5.11	4.00								
	p3-vor	4.74	5.08	10	17	12	12			2.00 5.69	
	p3-währ.	4.81	4.67	20	9	10					
	p3-nach	3.93	4.00								
	p4-vor	5.82	3.78								
Ärger	p4-währ.	5.40	3.56								
	p4-nach	4.64	3.69								
	BL	3.65	3.50								
	p1-vor	2.79	2.75							1.08 0.37 2.77	
	p1-währ.	3.31	3.25	11	12	16			++		
	p1-nach	3.28	3.27	11	14	13			+++		
	p2-vor	3.10	4.17						+		
	p2-währ.	3.64	3.33	10	17	12				2.00 9.38 xx 5.54 x	
	p2-nach	3.18	5.00	12	13	6					
	p3-vor	2.93	3.08	8	27	x	9				
	p3-währ.	3.30	3.50	7		13					
	p3-nach	3.37	3.33								
	p4-vor	3.82	3.78								
	p4-währ.	4.30	4.11								
	p4-nach	4.36	5.11								

Die routinemäßig mit t-tests berechneten Mittelwertsunterschiede zwischen Vor- und Hauptdiagnosen ergaben 2 signifikante Unterschiede.

Während die Verlaufspolygone der Gesamtmittelwerte für die Analogskalen Angst, Lampenfieber, Depression und Hilflosigkeit - und als gegenläufige Emotion: Freude - für die Datenerhebungszeitpunkte P1, P2 und P3 beeindruckend ausfallen, ist dies beim Zeitpunkt P4 weniger der Fall.

Statistisch untersuchten wir die Verlaufsdifferenzwerte "während" minus "vor" und "nach" minus "während" pro Analogskala pro Datenerhebungszeitpunkt. Wir testeten die empirischen Werte (Analogskalen gefallen, gleich geblieben, gestiegen) gegen die Nullhypothese der Zufallsvariation mit dem Chi-Quadrat-Test. In Tabelle 17 ist unter "Summe" der Gesamt-Chi-Quadrat-Wert angegeben, der mit $df=2$ als xxx = 0,1-%-Niveau, xx = 1-%-Niveau und x = 5-%-Niveau und (die Güte der Anpassung an die Zufallsverteilung) mit + = $p \leq .30$, ++ = $p \leq .50$ und +++ = $p \leq .70$ (als mäßig, gut und sehr gut) beurteilt wurde (vgl. CLAUSS-EBNER, S. 195 ff.). Zusätzlich beurteilten wir die Chi-Quadrat-Einzelwerte bei $df=1$, um einen heuristischen Hinweis auf die Bedeutung des Beitrags der einzelnen Zelle zu erhalten (vgl. KRAUTH & LIENERT, 1973, S. 28 ff.).

Bei Tabelle 17 ist zu beachten, daß bei der Differenzbildung Missing Values (vgl. S. 241) zu beachten waren und daß auf die Berechnung für den Zeitpunkt P4 wegen der geringen Stichprobengröße ($N=19$) verzichtet wurde.

Bei der Analogskala Angst finden wir signifikante Werte überwiegend für die Differenz der Skalierung der Angst während der Prüfung minus die Angst vor

der Prüfung. Überzufällig häufig ist der Wert kleiner als Null (i. e. die Angst fällt ab) und eher überzufällig selten größer als Null (i. e. die Angst steigt).

Die Analogskala "Lampenfieber" zeigt einen ähnlichen Verlauf, nur sind hier mehr die Differenzen "nach" minus "während" kleiner als Null.

Die Analogskala "Depression" ergibt - dort wo sie signifikant ausfallen - ähnliche Ergebnisse.

Die Analogskala "Freude" erreicht signifikante Werte beim Vergleich "während"/"nach"; die Freude steigt nach Abschluß der Prüfung.

Zu den signifikanten Ergebnissen der Analogskala "Wut", aber auch der Skala "Ärger", tragen die gleichbleibenden Skalierungen am meisten bei.

Die Kombination der beiden Differenzwerte - Differenz "während" minus "vor" und "nach" minus "während" - überprüften wir für jeden Datenerhebungszeitpunkt und jede Analogskala einmal bei Dichotomisierung nach dem Wert 0.0 ein andermal bei Trichotomisierung ($-0.5, 0.5$) durch Zweier-KFAn und - unter Einbeziehung der Prüfungsart - mit Dreier-KFAn. In keiner der 112 KFAn überlebten die Differenzfrequenzen den alpha-adjustierten Test gegen die Nullhypothese.

Die konfigurationsfrequenzanalytische Überprüfung der Differenzwerte der Datenerhebungszeitpunkte P1, P2 und P3 mit $N=39$ pro Analogskala durch Sechser-

KFAn und - unter Einbeziehung der Prüfungsart - durch Siebener-KFAn sowohl mit dichotomisierten als auch trichotomisierten (s. o.) Daten ergab beim alpha-adjustierten Test gegen die Nullhypothese ebenfalls keine einzige überzufällige Zellfrequenz.

Bei diesem Ergebnis ist allerdings zu beachten, daß die erwarteten Häufigkeiten jetzt nicht mehr über die theoretische Zufallsverteilung mit $p=.33$ (bzw. $p=.50$) berechnet werden, sondern über die empirischen Randverteilungen berechnet werden, z. B. $p(\text{Angst}/P1 \text{ w-v} / < 0.0) = 25/39 = .64$! Dadurch wird die erwartete Häufigkeit E für die Zellfrequenz ($\text{Angst}/PA/ --$) nicht mit $E=.33 \times .33 \times 39 = 4.33$, sondern mit $E=.64 \times .64 \times 39 = 16.03$ berechnet und fällt entsprechend höher aus. Würden wir die KFAn über den Chi-Quadrat-Ansatz berechnen, würde dies einen Unterschied von $\chi^2 (p=.33) = 98.56$ und $\chi^2 (p=.64) = 5.03$ bedeuten. Bei einer heuristischen KFA mit völliger Alpha-Adjustierung wäre der erste Wert signifikant, der zweite nicht. Wichtiges Ergebnis der KFAn, die wir oben erwähnten, ist, daß keine Differenzkonfigurationen, die eine Interaktion bedeuten würde (z. B. -O oder -+), signifikant ausfiel.

Auf der Basis der Rohwerte berechneten wir die KFAn für die Verläufe innerhalb der Datenerhebungszeitpunkte. Dichotomisiert wurde nach dem Median der Einzelmessungen (also z. B. dem Median aller Angstvor-Skalen). Die Mediane sind in Tabelle 18 wiedergegeben.

Tabelle 18: Mediane der Analogskalen

Analogskala	Meßzeitpunkt		
	"vor"	"während"	nach
Angst	6.5	4.5	3.1
Wut	2.0	2.0	2.5
Depress. Zustand	5.0	3.4	2.5
Lampenfieber	6.7	6.3	4.5
Freude	2.3	2.7	4.5
Hilflosigkeit	5.4	4.8	4.5
Arger	2.5	2.9	3.1

In Tabelle 19 sind die Ergebnisse von Dreier-KFAn für jede Analogskala und jeden Datenerhebungszeitpunkt wiedergegeben. Wegen der MV gehen unterschiedlich große Stichproben (=N) in die Analyse ein. Bei den signifikanten Typen bedeutet die Stufe 1 jeweils einen Wert niedriger als der Median.

Wie wir sehen, finden wir (mit einer Ausnahme) nur Kollektive von Merkmalsträgern, die entweder nur überdurchschnittliche oder nur unterdurchschnittliche Emotionsausprägungen skalierten. Zu beachten ist, daß in Tabelle 19 für jede verrechnete Variable ein neuer Median verwendet wurde.

So können wir zum Beispiel bei der Angst-P1-Messung insgesamt 57,9 % aller Versuchspersonen einem von zwei Typen zuordnen: dem Typ mit nur durchschnittlichen Merkmalsausprägungen (111) oder dem Gegentyp mit nur überdurchschnittlichen Merkmalsausprägungen (222).

Tabelle 19: Dreier-KFAn für die vor-während-nach-Messungen (v-w-n) jeder Analogskala für jeden Datenerhebungszeitpunkt getrennt; Missing Values wurden berücksichtigt; Mediane aus Tabelle 18

Variable	N	signif. Typ	F	E	S	s.	Q	s.
Angst P1 v-w-n	38	111 222	10 12	3,80 5,73	.340-02 .800-02	x -	.1787 xx .1014 xx	
P2 v-w-n	31	-						
P3 v-w-n	39	-						
P4 v-w-n	19	-						
Lampen- fieber P1 v-w-n	38	-						
P2 v-w-n	31	-						
P3 v-w-n	39	221	10	3,94	.450-02	x	.1727 xx	
P4 v-w-n	19	-						
Depr. P1 v-w-n	38	111 222	10 16	3,61 6,08	.230-02 .120-03	x xxx	.1831 xxx .3061 xxx	
P2 v-w-n	31	222	12	5,20	.320-02	x	.270 xx	
P3 v-w-n	39	111	12	6,00	.120-01	-	.1919 x	
P4 v-w-n	19	-						
Hilflo- sigkeit P1 v-w-n	39	-						
P2 v-w-n	3	222	15	7,98	.560-02	x	.2597 x	
P3 v-w-n	39	-						
P4 v-w-n	19	222	6	2,04	.110-01	-	.2344 xx	
Wut P1 v-w-n	39	222	7	1,57	.120-02	xx	.1448 xxx	
P2 v-w-n	31	111 222	15 9	6,32 2,15	.460-03 .180-03	xx xx	.3026 xxx .2086 xxx	
P3 v-w-n	39	222	8	2,07	.880-03	xx	.1606 xxx	
P4 v-w-n	19	111	5	1,33	.850-02	-	.2077 xx	
Ärger P1 v-w-n	38	-						
P2 v-w-n	31	111 222	11 10	3,82 3,87	.790-03 .330-02	xx x	.2302 xxx .1970 xx	
P3 v-w-n	39	111 222	14 13	4,95 4,69	.190-03 .430-03	xx xx	.2659 xxx .2421 xxx	
P4 v-w-n	19	-						
Freude P1 v-w-n	38	-						
P2 v-w-n	31	-						
P3 v-w-n	39	-						
P4 v-w-n	19	-						

Dieses Ergebnis können wir im Sinne einer linearen Korrelation interpretieren: Die gefundenen Ergebnisse gelten für die gesamte Skala. Dabei stört es nicht, daß der Gegentyp nur auf dem 10-%-Niveau signifikant ist: Der Test gegen die Nullhypothese entscheidet über die Berechtigung, von einem signifikanten Zusammenhang zu sprechen; für die Art des Zusammenhangs bzw. die Interpretation des Typs muß zusätzlich der jeweilige Gegentyp herangezogen werden.

Für Tabelle 19 gilt, daß außer für 'Lampenfieber-P3' keinerlei kurvilineare Zusammenhänge gefunden werden.

Die von uns hypostasierten Unterschiede zwischen Unerfahrenen und Erfahrenen überprüften wir, indem wir obige KFA_n unter Hinzufügen des Merkmals Prüfungsart (VD=1, HD=2) als Vierer-KFA_n wiederholten. Die Ergebnisse sind in Tabelle 20 dargestellt.

Wie zu sehen ist, findet sich auch hier keinerlei Interaktion.

Unterschiede zwischen den Prüfungsgruppen - etwa, daß die einen den durchschnittlichen Typ, die anderen den Gegentyp besetzen - finden wir nicht.

Ein anderes Problem ergibt sich aus der Eigenart unserer Datenerhebung: Nur die Analogskalen "nach" werden in vivo erhoben; die Analogskalen "vor" und "während" werden retrospektiv nach Bekanntgabe der Zensur bearbeitet. Für uns gibt es nachzuweisen,

Tabelle 20: Vierer-KFAn für die vor-während-nach-Messungen (v-w-n) jeder Analogskala für jeden Datenerhebungszeitpunkt getrennt unter Hinzufügung der Variablen Prüfungsart (Prüf.art)

Variablen in der KFA	N	signif. Typ	F	E	B	s.	Q	s.
Angst P1 v-w-n/Prüf.art	38	-						
P2 v-w-n/Prüf.art	31	-						
P3 v-w-n/Prüf.art	39	-						
P4 v-w-n/Prüf.art	19	-						
Lamp-entf. P1 v-w-n/Prüf.art	38	-						
P2 v-w-n/Prüf.art	31	111 1	9	2.83	.14D-02	x	.1919	xxx
P3 v-w-n/Prüf.art	39	-						
P4 v-w-n/Prüf.art	19	-						
Depr. P1 v-w-n/Prüf.art	38	222 1	12	4.21	.57D-03	xx	.2273	xxx
P2 v-w-n/Prüf.art	31	222 1	11	3.63	.51D-03	xx	.2349	xxx
P3 v-w-n/Prüf.art	39	-						
P4 v-w-n/Prüf.art	19	-						
Hilf1. P1 v-w-n/Prüf.art	38	-						
P2 v-w-n/Prüf.art	31	222 1	10	4.15	.55D-02	-	.1897	xxx
P3 v-w-n/Prüf.art	39	-						
P4 v-w-n/Prüf.art	19	111 2	6	1.31	.13D-02	x	.2650	xxx
Wut P1 v-w-n/Prüf.art	38	222 1	5	1.15	.56D-02	-	.1030	xxx
P2 v-w-n/Prüf.art	31	111 1	12	4.38	.64D-03	x	.2490	xxx
		222 1	8	1.49	.81D-04	xx	.1944	xxx
P3 v-w-n/Prüf.art	39	222 2	4	0.54	.37D-02	-	.0977	xxx
P4 v-w-n/Prüf.art	19	-						
Ärger P1 v-w-n/Prüf.art	38	-						
P2 v-w-n/Prüf.art	31	111 1	11	2.65	.29D-04	xxx	.2582	xxx
		222 1	9	2.68	.93D-03	x	.1956	xxx
P3 v-w-n/Prüf.art	39	111 1	9	3.47	.58D-02	-	.1567	xxx
		222 1	10	3.25	.11D-02	x	.1888	xxx
P4 v-w-n/Prüf.art	19	-						
Freude P1 v-w-n/Prüf.art	38	-						
P2 v-w-n/Prüf.art	31	-						
P3 v-w-n/Prüf.art	39	-						
P4 v-w-n/Prüf.art	19	-						

daß die retrospektive Messung unmittelbar im Anschluß an eine Prüfung (nahezu) ebenso valide Daten liefert wie eine in-vivo-Messung.

Die Tatsache, daß fast alle vergleichbaren Arbeiten sich ebenso wie wir auf retrospektive Messungen stützen, enthebt uns nicht des Problems, diesen Nachweis zu führen.

Wir gehen davon aus, daß die Problematik der retrospektiven Messung dann vernachlässigt werden kann, wenn eine vorgefundene starke Assoziation von Daten mit der Reliabilität der Skalen begründet werden kann.

Wir wollen Tabelle 19 betrachten. Hier finden wir bei den Analogskalen Depression, Hilflosigkeit, Wut und Ärger signifikante Typen in zwei oder mehr Datenerhebungszeitpunkten. Dies kann eine Folge der geringen Merkmalsfluktuation sein bzw. der hohen Reliabilität dieser Skalen, oder aber ein Ergebnis nachträglicher Dissonanzreduktion, Erinnerungseinflusses, Umbewertung o. ä.

Um nun nachzuweisen, daß dies eine Folge der Reliabilität der Skalen ist, rechneten wir Neuner-KFAn über die v-w-n-Skalen der Datenerhebungszeitpunkte P1, P2 und P3! Wir verweisen darauf, daß zwischen den Datenerhebungszeitpunkten im Schnitt 14 Tage liegen! In Tabelle 21 sind die Ergebnisse (und der Vollständigkeit halber die Ergebnisse der Zehner-KFAn unter Einbeziehung der Variable Prüfungsart) wiedergegeben.

Tabelle 21: Neuner-KFAn der Analogskalen mit 3 Meßzeitpunkten (v-w-n) und 3 Datenerhebungszeitpunkten (P1 - P2 - P3); Mediane aus Tabelle 18; Missing Values wurden berücksichtigt; in der unteren Hälfte Zehner-KFAn mit der zusätzlichen Variable 'Prüfungsart'

Variablen in der KFA	N	signif. Typ	F	E	B	s.	Q	s.
Angst P1-P2-P3/v-w-n	30	-						
Lamp- P1-P2-P3/v-w-n	30	-						
Depr. P1-P2-P3/v-w-n	30	111 111 111 222 222 222	4 5	0,03 0,10	.23D-07 .50D-07	xxx xxx	.1152 .1425	xxx xxx
Hilf1. P1-P2-P3/v-w-n	30	111 111 111	3	0,03	.34D-05	xx	.0862	xxx
Wut P1-P2-P3/v-w-n	30	111 111 111 222 222 222	7 3	0,40 0,005	.12D-06 .17D-07	xxx xxx	.1936 .0868	xxx xxx
Ärger P1-P2-P3/v-w-n	30	111 111 111 122 111 111 222 222 222	3 3 5	0,07 0,06 0,04	.56D-04 .34D-04 .67D-05	x x xxx	.0850 .0853 .1439	xxx xxx xxx
Freude P1-P2-P3/v-w-n	30	-						
Angst (s.o.)/Prüf.art	30	-						
Lampv (s.o.)/Prüf.art	30	-						
Depr. (s.o.)/Prüf.art	30	111 111 111 2 222 222 222 1	2 4	0,01 0,07	.38D-04 .69D-06	x xxx	.0577 .1142	xxx xxx
Hilf1. (s.o.)/Prüf.art	30	-						
Wut (s.o.)/Prüf.art	30	111 111 111 1 222 221 221 1 222 222 222 1	6 2 2	0,26 0,01 0,004	.19D-06 .57D-04 .72D-05	xxx - xx	.1679 .0577 .0575	xxx xxx xxx
Ärger (s.o.)/Prüf.art	30	111 111 111 1 122 111 111 1 222 222 222 1	3 3 4	0,05 0,04 0,03	.18D-04 .11D-04 .22D-07	x x xxx	.0856 .0858 .1152	xxx xxx xxx
Freude(s.o.)/Prüf.art	30	-						

Es zeigt sich, daß ein unterdurchschnittlicher Typ und ein überdurchschnittlicher Gegentyp den Test gegen die Nullhypothese der totalen Unabhängigkeit in den Skalen Depression, Wut und Ärger standhalten. Die Beurteilung der Signifikanz erfolgte über die totale Alpha-Adjustierung. Depressivität, Wut und Ärger scheinen demnach im Zusammenhang mit Prüfungen hochreliable Emotionen zu sein, während Angst und Lampenfieber von Prüfung zu Prüfung fluktuieren.

Gleichzeitig geht aus Tabelle 17 hervor, daß Angst und Lampenfieber - innerhalb einer Prüfung - den vorhergesagten abfallenden Verlauf zeigen. Genau entgegengesetzt - ansteigend - verläuft die skalierte Freude.

4.3 Hypothesenprüfender Teil

4.3.1 Analyse der Leistungsdaten (Hypothesen 1 und 2)

Die Hypothese 1 lautete: "Kandidaten in Abschlußprüfungen erhalten im Schnitt bessere Zensuren als Kandidaten in Zwischenprüfungen". In Hypothese 2 haben wir eine vergleichsweise höhere Durchfallquote in Zwischenprüfungen vorhergesagt.

In die Analyse des Komplexes 'Leistungsbeurteilung' gingen folgende Items ein (in Klammern der Datenerhebungszeitpunkt):

- "Welche Zensur erwartest Du als Gesamtzensur?" (BL)
- "Wie wären Deine Gefühle bei einer Gesamtzensur '3'?" (BL: 0 = Freude, 5 = Gleichgültigkeit, 9 = starker Ärger)
- "Wie wären Deine Gefühle bei einer Gesamtzensur '1'?" (BL: 0 = extreme Freude, 9 = völlige Verblüffung)
- "Welche Zensur erwartest Du in der Mehrheit der mündlichen Einzelprüfungen?" (BL)
- "Welche Zensur hast Du bekommen?" (P1 - P4, F2)
- "Welche Zensur wirst Du bekommen?" (P1 - P4, AN)
- "Mit welchem Prozentsatz des geprüften Stoffes hast Du Dich bereits während des Studiums (= frühere Semester) beschäftigt?" (P1 - P4; AN)

Von besonderem Interesse sind dabei die letzten drei der aufgeführten Variablen, da wir sie auf die Strategien und Belastungen beziehen werden.

Diese Items, in folgenden Leistungsdaten genannt, setzen wir außerdem in Beziehung zum fremdskalierten Item des Beisitzerblattes, ob der Beisitzer in seiner Zensureinschätzung von der des Prüfers - vor der Zensurendiskussion - abweicht.

In Tabelle 22 sind die erhaltenen Mittelwerte, Standardabweichungen, ggf. das Signifikanzniveau der t-tests für unabhängige Stichproben und des F-Tests für heterogene Varianzen wiedergegeben.

Tabelle 22: Mittelwerte und Standardabweichungen der Leistungsdaten

Item	Datenech. zeitpkt.	M(VD)	M(HD)	p(t-test)	s(VD)	s(HD)	p(F-Test)
Erwartete Gesamtzensur	BL	2,93	2,25	xx	0,53	0,45	
Gefühl bei Gesamtzensur "3" (0=Freude, 9=Ärger)	BL	3,76	6,42	xx	2,23	2,02	
Gefühl bei Gesamtzensur "1" (0=Freude, 9=Ver- blüffung)	BL	3,59	3,33		3,13	2,02	
M erwarteter Einzel- zensuren	BL	2,97	2,33	xx	0,50	0,50	
Erhaltene Zensur	P1	2,24	1,67		1,09	0,89	
	P2	2,24	2,00		1,02	1,13	
	P3	1,78	1,92		0,93	0,90	
	P4	2,54	1,56	x	1,13	0,53	x
Erwartete Zensur	P1	3,00	2,09	xxx	1,21	1,03	
	P2	3,07	3,00		0,83	0,92	
	P3	2,74	2,33		0,96	0,89	
	P4	3,27	2,33	x	1,08	0,98	
Prozentsatz des gepr. Stoffes im Studium	P1	44,14	60,82		28,61	28,54	
	P2	42,50	56,00		23,41	27,63	
	P3	48,89	64,08		25,22	26,10	
	P4	49,55	51,67		24,83	27,22	

Es zeigt sich, daß HDn ca. 14 Tage vor der ersten Prüfung nicht nur eine hochsignifikant bessere Gesamtzensur erwarten, sondern im Schnitt auch hochsignifikant bessere Einzelzensuren. Dementsprechend fallen auch die unterschiedlichen fiktiven Reaktionen auf eine 'befriedigende' Gesamtzensur aus: VDn freuen sich, HDn ärgern sich. Die Reaktionen auf eine Gesamt-Eins jedoch gleichen sich in beiden Gruppen: im Schnitt mittlere Freude.

Die Erwartung besserer Zensuren durch HDn setzt sich während der Prüfungszeit fort. In den Datenerhebungszeitpunkten P1 und P4 fielen die routinemäßig vorgenommenen t-tests signifikant aus, in den beiden anderen Datenerhebungszeitpunkten liegt der Schnitt der HDn niedriger als der der VDn.

Anders bei den tatsächlich erhaltenen Zensuren! Der Gesamt-Mittelwert liegt bei $M = 2.04$ mit $s = 1.01$ und $N = 140$; der Gesamtmittelwert für VDn beträgt $M = 2.14$ mit $s = 1.05$ und $N = 95$, für HDn $M = 1.82$ mit $s = 0.89$ und $N = 45$. Unabhängigkeit der Messungen vorausgesetzt, wäre der Mittelwertsunterschied nicht-signifikant.

Auf der Basis unserer $N = 140$ Messungen berechneten wir die Wahrscheinlichkeit (p), eine bestimmte Zensur zu erreichen:

Tabelle 23: Zensurwahrscheinlichkeiten auf der Basis von $N = 140$ Messungen

$p^{(1)}$	$= .357$
$p^{(2)}$	$= .371$
$p^{(3)}$	$= .185$
$p^{(4)}$	$= .071$
$p^{(5)}$	$= .021$

Die Wahrscheinlichkeit durchzufallen ist demnach geringer als die Wahrscheinlichkeit, beim Münzwurf hintereinander fünfmal 'Zahl' zu bekommen. Und die Wahrscheinlichkeit, mit 'sehr gut' oder 'gut' zensuriert zu werden, betrug mehr als 70 %. Die Chancen für die Kandidaten, eine möglichst gute Zensur zu bekommen, sind somit ausgezeichnet.

Bei 140 Messungen wurde ganze dreimal (2 VDN, 1 HD) eine '5' vergeben.

Wegen dieser geringen Fallzahlen konnte Hypothese 2 (Die Durchfallquote in Zwischenprüfungen ist höher) nicht geprüft werden.

Ebenso wie die Analyse der erhaltenen Zensur fällt die Analyse des Items 'Prozentsatz des Stoffes' aus: VDN skalieren im Schnitt geringere Werte als HDN, der univariate Unterschied ist jedoch nicht signifikant. Insgesamt werden im Schnitt 51,7 % des Prüfungsstoffes als während des Studiums bearbeitet skaliert. Traut man dieser Angabe, dann erarbeiten sich die Kandidaten rund die Hälfte des tatsächlich geprüften Stoffes kurzfristig vor den Prüfungen.

Einflüsse des Prüfers überprüften wir mit einfaktoriellen Varianzanalysen, wobei wir die $N = 140$ Messungen als unabhängig ansahen. Der Faktor Prüfer erwies sich bei den Variablen 'erwartete' und 'erhaltene Zensur' als hochsignifikant: die Rangfolge der Prüfer ist identisch. Die erwarteten Zensuren sind also bei dem Prüfer am höchsten, der auch die höchsten Zensuren gibt.

Aber: Die von uns berechenbaren Kontraste zwischen Prüfern des gleichen Prüfungsfaches fielen nicht-signifikant aus!

Die multivariaten Analysen führten wir wieder konfigurationsfrequenzanalytisch durch. Zugrunde legten wir die Gesamtmediane aller Messungen!

In Tabelle 24 sind die Ergebnisse wiedergegeben. Die KFAN, in denen die Variable 'Prüfer' hinzugefügt wurde, sind über alle vorhandenen Messungen gerechnet worden. Zu beachten ist, daß die Variablen 'erwartete Zensur' und 'Prozentsatz des geprüften Stoffes' dem Fragebogen AN entstammen; AN wurde von einigen Versuchspersonen unter ungeklärten Umständen bearbeitet, und diese Daten wurden daher aus den Analysen ausgeschlossen; die angegebenen Stichprobengrößen sind die tatsächlich verrechneten Fallzahlen.

Tabelle 24: Ergebnisse der KFA mit den Leistungsdaten; Missing Values wurden berücksichtigt

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. Typ	F	E	B	s.	Q	s.
Erwartete Zensur/ Prozentsatz/ Erhaltene Zensur	P1 P2 P3 P4	40 34 39 19	121 - 121 212	7 - 12 13	2,40 - 4,22 5,47	.900-02 - .610-03 .180-02	- xx x	.1208 xxx .2236 .2246	xx xxx xxx
-----P1 - P4		131	121 212	22 50	7,98 28,81	.140-04 .200-04	xxx xxx	.1100 .1986	xxx xxx
Erwartete Zensur/ Prozentsatz/ Erhaltene Zensur/Prüfungsart	P1 P2 P3 P4	40 34 39 19	1212 2121 - 2121	4 14 - 10	0,70 6,25 - 3,79	.520-02 .210-02 - .340-02	- x - -	.0829 .2263 xxx .1765	xxx xxx xxx xxx
-----P1 - P4		131	2121 1212	42 10	19,55 2,56	.600-07 .280-03	xxx xx	.1936 .0559	xxx xxx
Erwartete Zensur/ Prozentsatz/ Erhaltene Zensur/Prüfer	P1 P2 P3 P4	40 34 39 19	1212 2121 - 2121	18 15 - 5	8,64 5,14 - 2,73	.250-02 .210-03 - .00-00	- xx xxx xxx	.0714 .0756 xxx .0756	xx xxx xxx xxx
-----P1 - P4		131	2121 2123 1215	18 15 5	8,64 5,14 2,73	.250-02 .210-03 - .00-00	- xx xxx xxx	.0714 .0756 xxx .0756	xx xxx xxx xxx
Erwartete Zensur/ Prozentsatz/ Erhaltene Zensur/ Prüfer/Prüfungsart	P1 P2 P3 P4	40 34 39 19	1212 2121 - 2121	18 15 - 5	8,64 5,14 - 2,73	.250-02 .210-03 - .00-00	- xx xxx xxx	.0714 .0756 xxx .0756	xx xxx xxx xxx
-----P1 - P4		131	12152 21231	9 14	0,88 3,49	.300-06 .120-04	xxx xxx	.0603 .0796	xxx xxx

Die Analyse über alle N = 131 vollständigen Messungen ergibt in der Dreier-KFA der Variablen 'erhaltene und erwartete Zensur' und dem skalierten Prozentsatz des während des Studiiums bekanntgewordenen Stoffes zwei höchstsignifikante und prägnante Typen: eine Merkmalskombination mit überdurchschnittlichen erwarteten und erhaltenen Zensuren und überdurchschnittlichem Prozentsatz sowie

dem Gegentypen mit unterdurchschnittlich erhaltenen und erwarteten Zensuren nebst unterdurchschnittlichem Prozentsatz. Die Aufspaltung in die einzelnen Datenerhebungszeitpunkte ergibt, daß der Typ mit den überdurchschnittlichen Werten zu 86 % aus den Datenerhebungszeitpunkten P1 und P3 stammt. Wir gehen deshalb davon aus, daß das Problem der in der Gesamtanalyse mehrfach verrechneten Minikollektive bei diesem Datensatz vernachlässigt werden kann.

Auf das Problem mehrfach verrechneter Minikollektive werden wir an anderer (anschaulicherer) Stelle ausführlicher eingehen. Es geht kurz um folgendes: wenn man - wie wir oben im Lauf über $N=131$ - nicht zwischen abhängigen und unabhängigen Messungen unterscheidet, dann kann ein signifikanter Typ in der KFA allein deshalb auftreten, weil die Daten eines Minikollektivs von z. B. zwei Personen, bei denen die verrechneten Merkmalsausprägungen assoziiert sind, beispielsweise vier Mal erhoben wurden und die entsprechende Konfiguration dann acht Mal (2 Personen $\times$ 4 Messungen) in die KFA eingehen; die Interpretation des Typs als Assoziation von Merkmalsausprägungen müßte sich dann auf dieses Minikollektiv beschränken.

Die Hinzufügung des Merkmals Prüfungsart ($VD = 1$, $HD = 2$) ergibt bei der Gesamtanalyse wieder zwei hochsignifikante Typen: ein 'überdurchschnittlicher' Typ, der den HDn zugeordnet wird, und ein 'unterdurchschnittlicher' Typ, der den VDn zugeordnet wird. In das Vertrauensintervall des 'überdurchschnittlichen Hauptdiplomanden'-Typs fällt jedoch die Konfigurationsfrequenz der 'überdurchschnittlichen Vordiplomanden'; es ist also nicht zulässig, vom Merkmal Prüfungsart auf überdurchschnittliche erhaltene, erwartete und begründete Leistungen zu schließen.

Letztere Schlußfolgerung wird auch plausibel, wenn man beachtet, daß die Konfigurationsfrequenz der Dreier-KFA von $F=22$ auf die Konfigurationsfrequenz $F=10$ in der Vierer-KFA zurückgeht und sich gleichzeitig der Prägnanzkoeffizient fast halbiert.

Bei den 'unterdurchschnittlichen' Vordiplomanden fällt die Konfigurationsfrequenz des Gegentyps der unterdurchschnittlichen HDn nicht in das Vertrauensintervall des Typs und ist zudem nicht signifikant. Die Konfigurationsfrequenz $F=50$ in der Dreier-KFA geht nur auf $F=42$ in der Vierer-KFA zurück, der Prägnanzkoeffizient bleibt annähernd gleich. Wir haben also ein Kollektiv vor uns, in dem die Merkmale VD, erwartete Zensur '3' oder schlechter, erhaltene Zensur '2' oder schlechter und 'weniger als 50 % des Prüfungsstoffes im Studium begegnet' überzufällig häufig gemeinsam auftreten.

Für die Anthac-Analyse hatten wir Hypothese 1 verbal so formuliert: "Kandidaten in Abschlußprüfungen erhalten im Schnitt bessere Zensuren als Kandidaten in Zwischenprüfungen". In formalem Sinne haben wir damit zwei Typen hypostasiert: einen Typ, in dem die Merkmalsausprägung VD mit unterdurchschnittlichen Merkmalsausprägungen der Leistungsdaten assoziiert ist, und einen Typ, in dem die Merkmalsausprägung HD mit überdurchschnittlichen Merkmalsausprägungen assoziiert ist. Beide Typen haben wir als signifikant berechnet und können Hypothese 1 annehmen. Durch Konfigurationsvergleiche gelangen wir aber zu einer differenzierteren (Posthoc-) Aussage: Das 'Privileg' überdurchschnittlicher Leistungsdaten müssen sich die HDn mit einer kleinen Gruppe von VDn teilen, während unterdurchschnittliche Merkmalsausprägungen nur mit dem Merkmal 'VD' assoziiert ist.

In einem weiteren Lauf verrechneten wir die beiden Zensurvariablen und das Item des Prozentsatzes mit einer Nominalskala für die Prüfer. Wir erhielten drei überzufällige Typen. Das Kollektiv der Unterdurchschnittlichen rekrutiert sich zu zwei Dritteln aus Messungen bei den Prüfern 1 und 3, das Kollektiv der Überdurchschnittlichen zu zwei Dritteln aus Messungen beim Prüfer 1, 60 % aller Messungen beim Prüfer 3 und 31,3 % aller Messungen beim Prüfer 5 zugeordnet.

In einem letzten Lauf verrechneten wir die drei Leistungsvariablen, den Prüfercode und die Variable Prüfungsart. Wir erhielten zwei signifikante Typen: ein Kollektiv überdurchschnittlicher HDn beim Prüfer 5 und ein Kollektiv unterdurchschnittlicher VDn beim Prüfer 3. Dabei ist bemerkenswert, daß das Kollektiv der Überdurchschnittlichen HDn zu 90 % aus Kandidaten besteht, die Prüfungen beim Prüfer 5 ablegten. Das Kollektiv der unterdurchschnittlichen VDn verteilt sich jedoch gleichmäßig auf alle Prüfer, wie Konfigurationsvergleiche ergeben, erreicht für den Prüfer 3 jedoch das alpha-adjustierte Signifikanzniveau.

Man darf hierbei aber nicht vergessen, daß die Variable 'Prüfer' eng mit der Variable 'Prüfungsfach' assoziiert ist und daß alle berechenbaren Kontraste zwischen Prüfern desselben Prüfungsfaches nicht-signifikant ausfielen.

Zur Bestimmung des Einflusses des Prüfungsfachs zählten wir deshalb in einem ersten Schritt aus, wieviel Lehrveranstaltungen in den letzten zwei Semestern vor den Prüfungen in den einzelnen Prüfungsfächern stattfanden. In einem zweiten Schritt zählten wir aus, wieviel lernintensive Seminare oder Übungen in den vorhergehenden zwei Semestern

pro Prüfungsfach veranstaltet wurden. Den tatsächlichen Besuch der Lehrveranstaltungen ließen wir dabei unberücksichtigt und beschränkten uns allein auf die Auszählung des Lehrangebots. Dabei zeigte sich, daß die Assoziation der 'unterdurchschnittlichen VDN' mit dem Prüfer 3 einhergeht mit einer geringen Frequenz von Seminaren oder Übungen, die Assoziation der 'überdurchschnittlichen HDN' mit einer hohen Frequenz von Seminaren oder Übungen zu den von diesen Prüfern geprüften Fächern.

Wir analysierten die Beziehungen zwischen den 3 'Leistungsvariablen' auch parametrisch. In Tabelle 25 sind sowohl für alle N = 131 Messungen als auch für die einzelnen Prüfungszeitpunkte die Produkt-Moment-Korrelationen sowie die multiple Korrelation aufgeführt. Wie sich zeigt, ist die multiple Korrelation in keinem Fall wesentlich höher als die einfache Korrelation zwischen der erwarteten und der erhaltenen Zensur. Die Korrelationen des skalierten Prüfungsstoffes sind durchweg schwach oder nicht-signifikant.

Tabelle 25: Korrelationen der Leistungsdaten

Variable	P1 - P4		P1		P2		P3		P4	
	Proz. satz	Erh. Zens.	Proz. satz	Erh. Zens.	Proz. satz	Erh. Zens.	Proz. satz	Erh. Zens.	Proz. satz	Erh. Zens.
Erwartete Zensur	-.23	.64	-.34	.62	-.09	.98	-.29	.68	.04	.69
Erhaltene Zensur	-.16		-.18		-.11		-.24		-.16	
Multiple r	.635		.589		.581		.680		.720	
r ²	.404		.346		.337		.470		.520	

Hier zeigt sich deutlich, daß Korrelationen (i. e. eine mathematische Modellvorstellung mit linearen

und additiven Beziehungen, die für die gesamte Skala gelten,) und die Ergebnisse von Konfigurationsfrequenzanalysen (die Assoziationen und Interaktionen höherer Ordnung abbilden können, zudem Partialassoziationen und multiplikative Beziehungen zulassen,) verschiedene Modellvorstellungen sind.

Würden wir die Ergebnisse der Tabelle 25 unkritisch akzeptieren, dann müßten wir sagen, daß zwischen der Prüfungsleistung und den Studieninhalten nur eine geringe oder gar keine Beziehung besteht und daß die vom Kandidaten vorausgesagte und die tatsächlich erhaltene Zensur überraschend hoch miteinander korrelieren.

Die KFA hingegen errechnet, daß die jeweiligen überdurchschnittlichen Ausprägungen und die jeweiligen unterdurchschnittlichen Ausprägungen überzufällig assoziiert sind. Sie bildet damit auch die Korrelation von erwarteten und erhaltener Zensur ab.

In Tabelle 26 sind die Ergebnisse der gemeinsamen Analyse der Leistungsdaten, der Zensureinschätzung durch den Beisitzer dargestellt, sowie die eines Wiederholungslaufes unter Hinzufügung der Variablen 'Prüfungsart'.

Tabelle 26: Ergebnisse der KFA mit den Leistungsdaten und der Zensureinschätzung durch den Beisitzer (2 = weicht ab)

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. Typ	F	E	B	s.	Q	s.
Einsch.d.Beis.weicht ab erw.Zensur, Prozentsatz, erh.Zensur	P1 - P4	131	1121 1211 1212	21 4 42	6.37 13.42 25.65	.15D-o5 xxx .19D-o2 x .49D-o3 xx		.1174 xxx .o8o1 x .1552 xx	
Einsch.d.Beis.weicht ab erw.Zensur, Prozentsatz, erh.Zensur, Prüf.ert	P1 - P4	131	11212 12121	1o 35	1.75 18.6o	.12D-o4 xxx .13D-o3 xx		.o638 xxx .1459 xxx	

Insgesamt wurde das Item, ob die Beisitzer in ihrer Zensureinschätzung von der des Prüfers - vor Beginn der Zensurendiskussion - abweichen, in ca. 14 % der Prüfungen bejaht. Die Wahrscheinlichkeit einer Ja-Antwort ist annähernd gleich für Vor- und Hauptdiplomanden, für alle Prüfer und durch alle Beisitzer: alle Abweichungstests blieben nicht-signifikant.

Auffallend an den Ergebnissen der Tabelle 26 ist, daß keine der signifikanten Konfigurationen die Merkmalsausprägung der abweichenden Einschätzung (=2) enthält. Wir replizieren das 'Überdurchschnittliche' und das 'unterdurchschnittliche' Kollektiv, wobei beim 'Überdurchschnittlichen' kein Prägnanzverlust eintritt. Der Antityp besagt, daß die gleiche Einschätzung durch den Beisitzer überzufällig selten einhergeht mit unterdurchschnittlicher erwarteter, überdurchschnittlich erhaltener Zensur und unterdurchschnittlich bekanntem Prüfungsstoff.

Insgesamt finden wir, daß bei den Messungen, bei denen die vom Kandidaten erwartete und die erhaltene Zensur übereinstimmend über- oder unterdurchschnittlich ausfallen, überzufällig häufig der Beisitzer zur gleichen Bewertung gelangt wie der Prüfer.

Die Ergebnisse lassen sich zusammenfassen wie folgt:

- Die Wahrscheinlichkeit, in der Prüfung durchzufallen, ist im untersuchten Prüfungszeitraum geringer als die Wahrscheinlichkeit, beim Münzwurf hintereinander fünfmal 'Zahl' zu bekommen. Hypothese 2 (Die Durchfallquote in Zwischenprüfungen ist höher als in Abschlußprüfungen) konnte deshalb nicht geprüft werden; die theoretischen Überlegungen zur Selektionsfunktion von Prüfungen (die zur Formulierung von Hypothese 2 führten) konnten nicht bestätigt werden.
- Das Kollektiv der HDn ist mit überdurchschnittlichen Merkmalsausprägungen der Leistungsdaten assoziiert, muß dieses 'Privileg' aber mit einem kleinen Kollektiv von VDn teilen.
- Die Merkmalsausprägung VD der Variablen 'Prüfungsart' ist mit unterdurchschnittlichen Merkmalsausprägungen der Leistungsdaten assoziiert.
- Wir fanden eine starke Assoziation zwischen der nach der Prüfung (vor Bekanntgabe der tatsächlichen Zensur) vom Kandidaten geschätzten und erwarteten Zensur, der Skalierung der Menge des Prüfungsstoffes (der vor Beginn der Prüfungsvorbereitungen bekannt war) und der tatsächlich erhaltenen Zensur.

Versuchspersonen, die eine Überdurchschnittliche Zensur erwarten und erhalten, sowie einen Überdurchschnittlichen Prozentsatz skalieren, nennen wir fortan 'Überdurchschnittliche', den Gegentyp bezeichnen wir mit 'Unterdurchschnittliche'.

- Unterschiede zwischen Prüfungen bei verschiedenen Prüfern lassen sich auch darauf zurückführen, daß diese Prüfer Fächer prüften, für die in unterschiedlichem Ausmaß lernintensive Lehrveranstaltungen angeboten wurden.

4.3.2 Analyse der Belastungsmessungen

(Hypothesen 6 und 8)

Als sechste Hypothese hatten wir formuliert: "Die Indikatoren von Angst und Depression korrelieren hoch positiv und bilden gemeinsam die Symptomatik der Prüfungsbelastung". Hypothese 8 lautete: "Kandidaten in Zwischenprüfungen leiden unter stärkeren Prüfungsbelastungen als Kandidaten in Abschlußprüfungen".

In diesem Abschnitt werden wir die einzelnen Datensätze (Meßinstrumente) getrennt für sich z. B. auf faktorielle Validität oder auf Retestreliabilität untersuchen. In der zweiten Hälfte dieses Abschnitts werden wir die Datensätze dann agglutinieren und die statistischen Ergebnisse der Hypothesenprüfungen vorstellen.

4.3.2.1 Analyse der einzelnen Meßinstrumente

4.3.2.1.1 Die Hilflosigkeitsskala (HiS)

Abbildung 13 zeigt den Eigenwertverlauf einer R-Hauptkomponentenanalyse mit 20 HiS-Items der BL-Erhebung. Nach dem Scree-Test ist ein Faktor zu extrahieren. Auf diesem Faktor laden alle Items substantiell ($\alpha > |.40|$) und positiv. Die vorgesehene A-priori-Testscorebildung durch Summation über alle Items wird durch diese Ergebnisse bestätigt.

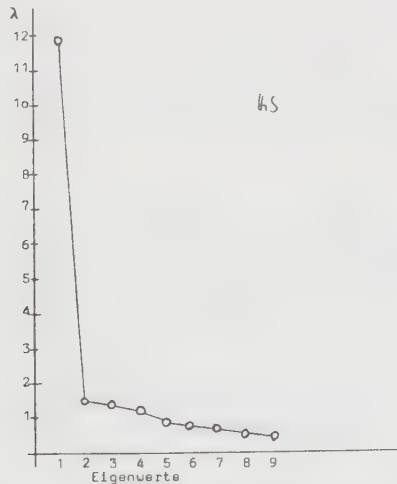


Abbildung 13: Eigenwertverlauf einer Hauptkomponentenanalyse der HiS-Items des Datenerhebungszeitpunktes BL

In Tabelle 27 sind die - auf der Basis der Testscores berechneten - statistischen Kennwerte wiedergegeben.

Tabelle 27: Ergebnisse der Messungen mit der HiS

Datenerhebungszeitpkt.	M	s	n(VD)	M(VD)	n(HD)	M(HD)	p(t-test)
BL	42.76	14.99	29	45.38	12	35.08	x
P1	48.42	14.51	29	48.62	12	47.92	
P2	55.00	16.77	29	57.07	12	50.17	
P3	48.03	14.72	27	47.48	12	49.25	
P4	52.03	15.72	10	54.73	9	48.78	
Gesamtmean	50.71	15.54	Gesamtmedian		Med= 49.75		

Der Mittelwertsverlauf ist in Abbildung 14 dargestellt.

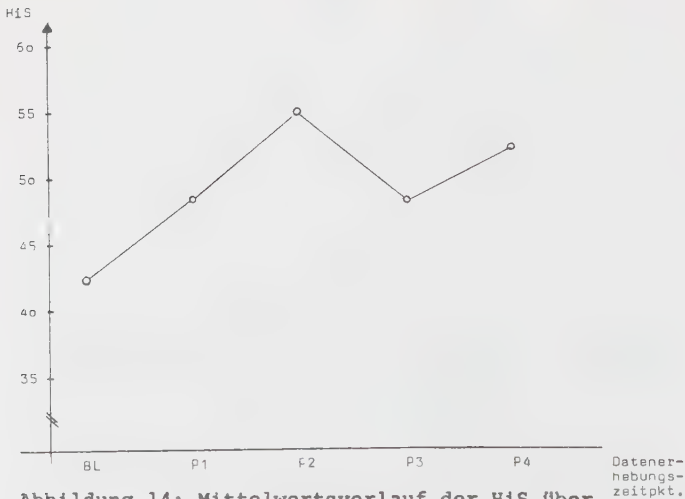


Abbildung 14: Mittelwertsverlauf der HiS über alle Datenerhebungszeitpunkte

Univariate Varianzanalysen zur globalen Einschätzung der Effekte von Datenerhebungszeitpunkten zum einen und des Faktors "Prüfer" zum anderen brachten keine signifikanten Ergebnisse.

Eine Vierer-KFA der gesamtmediandichotomisierten Testscores über die Datenerhebungszeitpunkte BL, P1, P2 und P3 sowie eine Wiederholung als Fünfer-KFA unter Einbeziehung der Variablen "Prüfungsart" ergab keinen signifikanten Typ.

In KFA mit den Verlaufsdifferenzwerten (also z. B. $HiS (D1) = HiS (P1) - HiS (BL)$ und $HiS (D2) = HiS (P2) - HiS (P1)$ usw. ergab sich ebenfalls keine überzufällige Konfiguration (Dichotomisierung nach $M = 0.0$).

Dies heißt zusammenfassend, daß die HiS-Scores allein in keinem systematischen (überzufälligen) Zusammenhang mit den Variablen Prüfer, Prüfungsart und Datenerhebungszeitpunkt stehen.

4.3.2.1.2 Die Selbstrating-Kurzform des Beck-Depression-Inventory (BDI)

In Abbildung 15 ist der Eigenwertverlauf einer Hauptkomponentenanalyse der gewichteten Daten des Datenerhebungszeitpunkts BL dargestellt. Er legitimiert die von uns vorgenommene einfache Summenscorebildung.

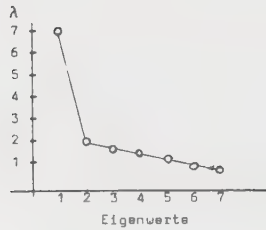


Abbildung 15: Eigenwertverlauf einer Hauptkomponentenanalyse der Items des Datenerhebungszeitpunktes BL

Abbildung 16 gibt den Verlauf der Werte wieder.

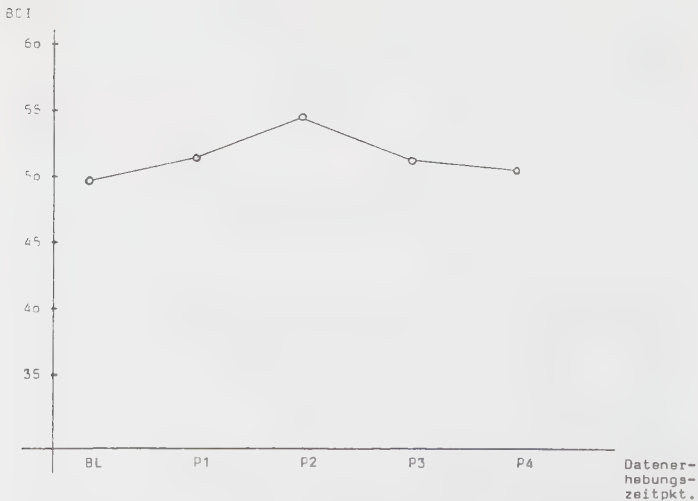


Abbildung 16: Mittelwertsverlauf der BDI-Scores über alle Datenerhebungszeitpunkte

In Tabelle 28 sind die dazugehörigen Statistiken wiedergegeben.

Tabelle 28: Ergebnisse der Messungen mit dem BDI

Datenerhebungszeitpkt.	M	s	M(VD)	M(HD)	p(t-test)
BL	49.75	13.41	51.83	44.75	
P1	52.39	11.39	53.93	48.67	
P2	54.33	13.34	56.14	50.08	
P3	51.93	11.51	51.96	51.92	
P4	50.95	9.97	53.27	48.11	
Gesamt Daten	52.61	11.77	Gesamtmedian Med= 51.70		

Einfaktorielle Varianzanalysen mit den Faktoren 'Prüfer' und 'Datenerhebungszeitpunkt' blieben ohne signifikante Ergebnisse.

KFAn mit den wiederholt gemessenen mediandichotomisierten Testscores (Fünfer-KFAn mit BL, P1, P2, P3, P4 bei N = 19; Vierer-KFAn mit BL, P1, P2, P3 bei N = 39; Dreier-KFAn mit BL, P1, P2 bei N = 41) sowie Wiederholungsläufe unter Einbeziehung der Variablen Prüfungsart (VD/HD) erbrachten keinen signifikanten Typ.

Überzufällige Konfigurationsfrequenzen fanden sich ebenfalls nicht bei den KFAn mit Verlaufsdifferenzwerten (M = 0.0).

Dies bedeutet zusammenfassend, daß die BDI-Scores allein in keinem systematischen Zusammenhang mit den Variablen 'Prüfer', 'Prüfungsart' und 'Datenerhebungszeitpunkt' stehen.

4.3.2.1.3 Das State-Trait-Anxiety-Inventory Form X1 (STAI)

In Abbildung 17 ist der Eigenwertverlauf der BL-Daten des STAI dargestellt.

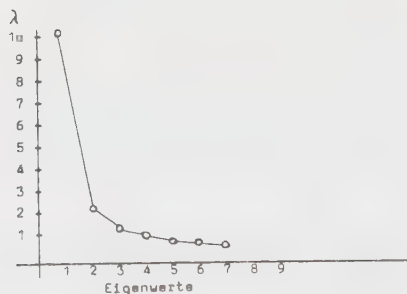


Abbildung 17: Eigenwertverlauf einer Hauptkomponentenanalyse der Items des STAI des Datenerhebungszeitpunktes BL

Entsprechend dem Scree-Test extrahieren wir zwei Faktoren. Die Items mit Ladungen größer als $a = |.40|$ sind in Tabelle 29 dargestellt.

Tabelle 29: Faktorladungen $a > |.40|$ der Items des STAI

Item Nr.	Text	F1	F2
17.	Ich mache mir Sorgen	.31	
6.	Ich bin aufgeregt	.64	
7.	Ich mache mir zur Zeit Sorgen, daß etwas schief gehen könnte	.84	
9.	Ich bin beunruhigt	.84	
12.	Ich fühle mich nervös	.80	
4.	Ich bedaure etwas	.59	
14.	Meine Nerven sind bis zum Zerreißen angespannt	.67	.45
13.	Ich bin zappelig	.62	
18.	Ich bin überreizt und durcheinander	.61	
11.	Ich habe Selbstvertrauen	-.43	-.43
15.	Ich bin entspannt	-.50	-.65
2.	Ich fühle mich sicher	-.65	-.42
1.	Ich fühle mich ruhig	-.69	
20.	Ich bin vergnügt		-.77
5.	Ich fühle mich wohl		-.77
19.	Ich fühle mich froh		-.77
10.	Ich fühle mich behaglich		-.80
16.	Ich fühle mich zufrieden		-.90
8.	Ich fühle mich ausgeruht		-.59

Faktor 1 (F1) ist bipolar; auf dem einen Pol laden Items wie Beunruhigung, Nervosität und Aufgeregtheit, auf dem anderen Pol Items der Ruhe und Entspannung. F1 interpretieren wir deshalb als Aktiviertheits-faktor. Faktor 2 (F2) ist (fast) unipolar und erfaßt die Items des Vergnügens, Wohlbefindens, Frohsinns und der Behaglichkeit. Wir interpretieren F2 als Valenzfaktor. Dieses Ergebnis würde bedeuten, daß die mit dem STAI gemessene 'Angst' sich aus den Dimensionen 'Aktiviertheit' und 'Valenz' zusammensetzt.

Andererseits ist es möglich, daß die zweifaktorielle Lösung ein Ergebnis eines Artefakts ist und nicht einer inhaltlichen Gruppierung. Die Häufigkeitsverteilungen der positiv formulierten Items, die mit einer Ausnahme alle substantielle Ladungen auf dem Valenzfaktor F2 haben, sind überwiegend rechtsschief; die der negativ formulierten Items, die - mit einer Ausnahme - alle nur auf dem Aktiviertheitsfaktor laden, sind überwiegend linksschief. Die zweifaktorielle Lösung kann u. U. deshalb auch eine Folge von Verteilungseigenheiten sein und etwas Ähnliches wie 'Methodenfaktoren' darstellen.

Wir standen vor der Entscheidung, entweder Faktorscores für die beiden Dimensionen zu berechnen oder die von SPIELBERGER, GORSUCH & LUSHENE (1970) angegebene Summensecorebildung zu übernehmen. Da wir uns auf MENGEN (1971, zitiert nach RACHMAN & BERGOLD, 1976) als Referenzarbeit berufen, da ein heterogener Testkernwert stabiler ist und da wir sowieso die Zahl der Variablen im Laufe der Analysen verringern müssen, wählten wir die Summensecorebildung.

Abbildung 18 zeigt den Verlauf des arithmetischen Mittels der Testscores über alle Datenerhebungszeitpunkte.

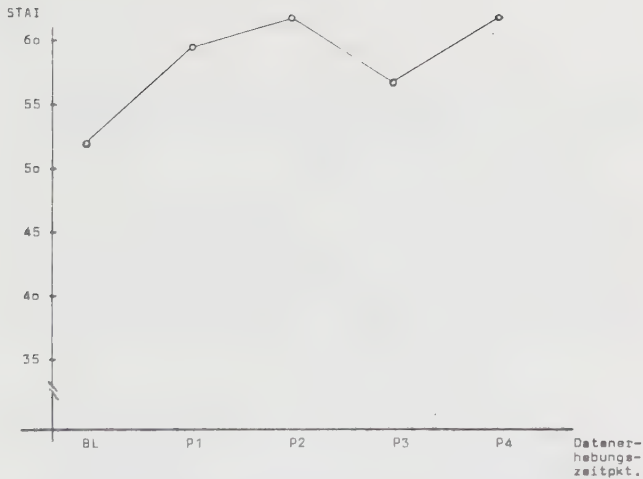


Abbildung 18: Mittelwertsverlauf der STAI-Scores über alle Datenerhebungspunkte

In Tabelle 30 sind die dazugehörigen statistischen Daten angegeben.

Tabelle 30: Ergebnisse der Messungen mit dem STAI

Datenerhebungszeitpkt.	M	s	M(VD)	M(HD)	p(t-test)
BL	51.76	14.39	52.55	49.83	
P1	59.27	14.01	59.61	59.62	
P2	61.53	13.26	62.96	59.17	
P3	56.07	14.00	56.07	57.00	
P4	61.40	12.47	63.45	58.83	
Gesamtdaten	59.41	13.61	Gesamtmedian Med= 58.70		

Einfaktorielle Varianzanalysen mit den Faktoren Prüfer und Datenerhebungszeitpunkt ergaben keine signifikanten Unterschiede.

Die nach dem Gesamtmedian dichotomisierten Daten verrechneten wir für die Datenerhebungszeitpunkte BL, P1, P2, P3 und P4 mit $N = 19$ in Fünfer-KFAn und - unter Hinzufügung der Variable Prüfungsart - in Sechser-KFAn; für die Datenerhebungszeitpunkte BL, P1, P2 und P3 in Vierer- (Fünfer-) KFAn.

Bei letzterwähnter Vierer-KFA erhielten wir einen signifikanten Typ 1111 mit $F = 9$, $E = 3.17$, $B = .34D-02$ und $Q = .1627$ (xx), der auf dem 10-%-Niveau signifikant war. Dies ist ein Hinweis auf ein Kollektiv, das konstant unterdurchschnittliche Angst skalierte.

Die Analyse der Verlaufsdifferenzwerte mit Vierer- (Fünfer-) KFAn für $N = 19$ und Dreier- (Vierer-) KFAn für $N = 39$ ergab keine überzufällig häufige oder seltene Konfigurationsfrequenz.

Faktorenanalytisch zeigt sich der STAI als heterogener Test. Erstmals in dieser Arbeit können wir ein Kollektiv mit stabilen emotionalen Reaktionen nachweisen.

4.3.2.1.4 Analyse der Beschwerdeliste

Um die physiologischen Reaktionen zu messen, hatten wir aus der Freiburger Beschwerdeliste 18 Items entnommen, sie mit den Antwortstufen 'ständig - fast immer - ein- bis zweimal - nie' versehen und sie F2 beigelegt mit der Frage: "Wie oft traten die folgenden Symptome während der Prüfung bei Dir auf?" Beim Datenerhebungszeitpunkt BL hatten wir auf eine Verwendung der Beschwerdeliste verzichtet, da keine akuten Prüfungsbelastungen zu messen waren.

In einem ersten Analyseschritt rechneten wir für jeden Datenerhebungszeitpunkt R-Hauptkomponentenanalysen mit anschließender Varimaxrotation. Die Eigenwertverläufe sind übereinanderprojiziert in Abbildung 19 dargestellt.

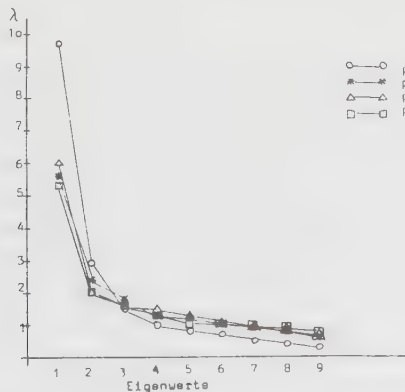


Abbildung 19: Übereinanderprojizierte Eigenwertverläufe der Hauptkomponentenlösungen für die Beschwerdelisten der einzelnen Datenerhebungszeitpunkte

Nach dem Scree-Test extrahierten wir für jeden Datenerhebungszeitpunkt zwei Faktoren, deren varimaxrotierte Ladungen in Tabelle 31 dargestellt sind. Ladungen, die kleiner als $a = |.40|$ waren, sind nicht dargestellt.

Tabelle 31: Faktorladungen und Kommunalitäten aus den Hauptkomponentenanalysen der Beschwerdelistenitems, ermittelt getrennt für jeden Prüfungs-Datenerhebungszeitpunkt

Item	Text	Faktorladungen								Kommunalitäten			
		F1				F2				P1	P2	P3	P4
		F1	F2	F2	F1	F2	F1	F2	F2				
2.	Überraschendes Ermüden	.80		.79		.78		.77		.74	.71	.60	.51
3.	Schwindelgefühle	.78		.57		.53		.48		.73	.45	.53	.44
4.	Müdigkeit u. Zerschlagenh.	.77		.70		.73		.62		.71	.55	.70	.46
5.	Brennbarkeit	.74		.54	.49		.62		.57	.54		.70	.70
6.	Gefühl, a.d. Toilette z.m.	.58		.13		.62		.96		.36	.53	.41	.73
7.	Kopfschmerzen	.56		.73		.43		.93		.33	.40	.33	.66
8.	er. Empfinden i. Nacken	.55			.76		.44		.93	.33	.40	.30	.72
9.	Gefühl, gleich Leinen z.m.	.50		.48		.50		.81	.73	.30	.70	.40	.40
10.	Unruhe	.45		.31		.51		.96		.20	.40	.45	.43
11.	Atembeschwerden	.43			.73		.50		.95	.20	.70	.41	.73
12.	Anfallsartig. Herzklopfen		.44		.46		.49		.74	.26	.70	.70	.53
13.	Krampf i. Bauchraum		.43		.44		.73		.74	.26	.70	.33	.56
14.	Feuchte Hände		.42		.50		.80		.84	.23	.63	.45	.71
15.	Stottern		.32		.63		.62		.80	.20	.69	.43	.56
16.	Schwitzen		.31		.70		.81		.86	.20	.69	.70	.76
17.	Zittern o. "weiche Knie"		.31		.68		.61	.40	.79	.20	.67	.43	.57
18.	Mitbewellen		.32		.79		.63	.50	.76	.20	.67	.45	.44
19.			.31		.76		.51		.86	.20	.73	.26	.44

Die vollständigen Faktorenlösungen verglichen wir nach dem Verfahren von FISCHER & ROPPERT (1964). Die Faktormatrix P1 war invariant. Die Ergebnisse sind in Tabelle 32 dargestellt.

Tabelle 32: Ähnlichkeitskoeffizienten des Faktormatrizenvergleichs der Beschwerdelisten nach dem Verfahren von FISCHER & ROPPERT (1964)

Datenerhebungszeitpkt.	P1	P2	P3	N
P2	.9286			41
P3	.9102	.9412		39
P4	.7627	.7753	.7155	19
N	41	41	39	140

Für die Ähnlichkeitskoeffizienten gibt es keine Signifikanztests. GEBHARDT (1968) nennt als untere Grenze 0.8: Koeffizienten unter 0.8 sollen nur mit großer Vorsicht als Hinweis auf die Ähnlichkeit der Faktorenstruktur interpretiert werden.

Danach können wir von ähnlichen Faktorstrukturen der Datenerhebungszeitpunkte P1, P2 und P3 sprechen. Die Ähnlichkeitskoeffizienten für P4 liegen unter dem von GEBHARDT genannten kritischen Wert, erreichen aber jeweils annähernd die gleiche Größenordnung nicht viel kleiner als 0.8. Zum Teil oder überwiegend dürfte dies auf die im Vergleich zur Stichprobengröße ($N = 19$) hohe Variablenzahl ($NVAR = 18$) zurückzuführen sein.

In allen vier Lösungen erreichen die Items 'überraschendes Ermüden', 'Schwindelgefühle', 'Müdigkeit' und 'Zerschlagenheit', 'Gefühl, auf die Toilette zu

müssen', 'Kopfschmerzen', 'Gefühl, gleich weinen zu müssen' und 'Übelkeit' substantielle Ladungen auf demselben einen Faktor. Dies sind Symptome des von SCHULTE & TOELLE beschriebenen "psychovegetativen Erschöpfungssyndroms"; wir wollen den Faktor deshalb 'Erschöpfungssyndrom' nennen.

Die Items 'Erröten', 'Hitzewellen', 'Zittern oder weiche Knie', 'Schwitzen', 'Stottern', 'feuchte Hände' und 'anfallsartiges Herzklopfen' erreichen substantielle Ladungen auf demselben einen Faktor, den wir als 'Angstsyndrom' interpretieren wollen.

Für jeden Prüfungstermin getrennt berechneten wir über den Regressionsansatz die Faktorscores.

Da beim Datenerhebungszeitpunkt P4 die Korrelationsmatrix singulär war, wurde die Variable 18 ('Übelkeit') bei der Scoreberechnung unberücksichtigt gelassen, weil sie zusammen mit Variable 5 ('Kopfschmerzen') eine extrem hohe Ladung auf dem einen Faktor ($\alpha = .96$ und $\alpha = .93$) und eine Nullladung auf dem anderen Faktor ($\alpha = .062$ und $\alpha = .005$) erreichte; die Faktorscores für P4 wurden dann auf der Basis von 17 Variablen berechnet. Die Ladungen auf den dazugehörigen varimaxrotierten Faktoren änderten sich durch die Reduktion um diese eine Variable um maximal .036, die Varianz des Erschöpfungsfaktors reduzierte sich um 2 %, die Varianz des Angstfaktors nahm um 0.8 % zu.

Die Faktorscores verwendeten wir als Variablen für Konfigurationsfrequenzanalysen. Zur Dichotomisierung verwendeten wir das bei der Faktorscoreberechnung verwendete Arithmetische Mittel; nach der Dichotomisierung bedeutet die Stufe '2' 'überdurchschnittliche Erschöpfung' und 'überdurchschnittliche Angst'.

Wir rechneten KFA mit den Daten der Zeitpunkte P1/P2/P3/P4 (N = 19), P1/P2/P3 (N = 39) und P1/P2 (N = 41) als Variablen. Wir rechneten sie einmal nur für die Erschöpfungsscores, einmal nur für die Angstscores und einmal multivariat mit beiden; alle Läufe wurden unter Hinzufügung der Variablen Prüfungsart wiederholt.

In den insgesamt 18 Läufen fanden wir nur einen einzigen signifikanten Typ: Bei der Dreier-KFA mit den Variablen P1-E, P2-E und Prüfungsart (VD = 1) erwies sich die Konfiguration 221 mit $F = 9$, $E = 3.38$, $B = .54D-02$ (x) und $Q = .1514$ (xx) als signifikant auf dem 5-%-Niveau; in dem Lauf ohne die Variable Prüfungsart mit den Variablen P1-E und P2-E ergab sich auf dem 10-%-Niveau ein Hinweis auf einen Typ 22 mit $F = 10$, $E = 4.78$, $B = .17D-01$ (-) und $Q = .1464$ (x).

Zusammenfassend läßt sich sagen, daß die konfigurationsfrequenzanalytische Überprüfung der Beschwerdelistenfaktorscores enttäuschende Ergebnisse brachte.

Während die Variablenstruktur - faktorenanalytisch - überzeugend reliabel ist, muß nach Betrachtungen der KFA-Ergebnisse geschlossen werden, daß die Messungen der einzelnen Meßwertträger nicht reliabel sind, i. e. das gemessene Merkmal in starkem Ausmaß fluktuiert.

4.3.2.1.5 Analyse der Analogskalen zur Erfassung der Nachprüfungsbefindlichkeit

Mit vier Analogskalen hatten wir versucht, den psychischen Zustand nach der Prüfung vor Bekanntgabe der Zensur zu erfassen. Es waren die Skalen

- "Gesetzt den Fall, der Prüfer würde Dir gleich die Chance geben, Deine Zensur durch eine Nachprüfung (5 Minuten) jetzt sofort zu verbessern. Wie würdest Du reagieren?"

(1 = spontan zustimmen, 9 = spontan ablehnen)

- "Wie erschöpft fühlst Du Dich im Moment?"

(1 = fühle mich voll fit, 9 = so erschöpft fühlte ich mich noch nie)

- "Wie schätzt Du Deine momentane Konzentrationsfähigkeit ein?"

(1 = so hoch wie noch nie, 9 = so niedrig wie noch nie)

- "War der Verlauf der Prüfung genauso, wie Du Dir das vorgestellt hattest?"

(1 = nein, viel schlechter, 5 = ja, ganz genauso, 9 = nein, viel besser)

In Tabelle 33 sind die Mittelwerte und Standardabweichungen angegeben.

Tabelle 33: Befindlichkeit nach der Prüfung vor Bekanntgabe der Zensur

Item (Kurztext)	Datenerh. zeitpkt.	M	s	M(VD)	M(HD)	p(t-test)
Nachprüfung(9=abl.)	P1	3,95	2,35	4,38	2,92	xx
	P2	4,63	2,94	4,82	3,25	
	P3	4,51	3,00	5,41	2,50	
	P4	4,25	3,32	5,09	3,22	
Erschöpfung(9=wie nie)	P1	4,60	1,93	4,72	4,27	
	P2	5,47	1,93	5,50	5,25	
	P3	4,77	2,10	4,74	4,83	
	P4	5,40	2,26	5,71	5,00	
Konzentration(9=niedr.)	P1	4,37	1,75	4,45	4,18	
	P2	5,41	1,72	5,39	5,50	
	P3	5,33	1,00	5,33	5,33	
	P4	5,35	2,13	5,73	4,89	
Verlauf d.Pr.(5=wie erw., 9=besser)	P1	4,72	2,40	5,07	3,82	
	P2	4,38	2,44	4,46	3,75	
	P3	4,77	2,10	4,74	4,83	
	P4	3,95	1,36	4,45	3,33	

Es zeigt sich, daß die Mittelwerte um die Mittel-
punkte der Skalen liegen. Bedeutung hat dies bei
der Einschätzung des Prüfungsverlaufes, der im
Mittel als genauso wie erwartet skaliert wird.

Einfaktorielle univariate Varianzanalysen über
alle N = 131 (!) Messungen für den Faktor Prüfer
ergaben keine signifikanten Ergebnisse.

Mit Konfigurationsfrequenzanalysen für alle vier
Variablen unter Hinzufügung der Variable Prüfungs-
art ($VD = 1/HD = 2$) überprüften wir die multiva-
riaten Zusammenhänge. Die reduzierte Stichproben-
größe entsteht durch Missing Values. Wir rechneten

alle berechenbaren Zweier-, Dreier-, Vierer- und Fünfer-KFAn. In Tabelle 34 haben wir die Läufe mit signifikanten Ergebnissen zusammengestellt.

Alle Läufe für die Datenerhebungszeitpunkte P2, P3 und P4 blieben ohne signifikante Typen. Zugrunde gelegt wurden jeweils die Gesamtmediane; die Variable 'Verlauf der Prüfung' wurde trichotomisiert.

Auffallend ist das Ergebnis der Zweier-KFA der Variablen 'Erschöpfung' und 'Konzentration' mit $N = 131$. Die überzufällig häufigen Typen 11 und 22 und die überzufällig seltenen Typen 12 und 21 stellen die perfekte konfigurationsfrequenzanalytische Darstellung einer linearen positiven Korrelation dar: Überzufällig häufig finden wir Merkmalsträger mit gleichsinnigen Merkmalsausprägungen, überzufällig selten Merkmalsträger mit gegensinnigen Merkmalsausprägungen.

Die Hinzufügung des Merkmals Prüfungsart erhellt die Korrelation noch mehr. Das Kollektiv der Konzentrierten und Nicht-Erschöpften erreicht nicht mehr das alpha-adjustierte Signifikanzniveau, d. h. dieses Kollektiv ist mit dem Merkmal Prüfungsart nicht assoziiert.

Tabelle 34: KFAn mit signifikanten Ergebnissen bei den Variablen zur Messung der Nachprüfungsbefindlichkeit (nähere Angaben im Text)

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. typ	F	E	B	s.	Q	s.
Nachp., Erschöpfung	P1 - P4 P1 P2 P3 P4	131 40 33 39 19	- - - - -						
Nachp., Erschöpfung, Prüf.art (VD=1, HD=2)	P1 - P4 P1 P2 P3 P4	131 40 33 39 19	112 - - - -	20	10,28	.310-02	x	.0805	xx
Erschöpfung, Konzentration	P1 - P4 P1 P2 P3 P4	131 40 33 39 19	11 12 21 22 22 9 4,20	52 19 18 41 27,94	37,94 33,06 32,06 27,94	.540-02 .210-02 .180-02 .280-02 .210-01	x xx xx x -	.1511 .1436 .1421 .1363 .1341	x x x xx x
Erschöpfung, Konzentration, Prüf.art	P1 - P4 P1 P2 P3 P4	131 40 33 39 19	121 211 221 221 -	10 12 36 3 3,05	23,98 23,35 20,26 3,05	.440-03 .440-02 .310-02 .960-02	xx x xx -	.1306 .1044 .1421 .1341	xx x xxx xx
Nachpr., Erschöpfung, Konzentration	P1 - P4 P1 P2 P3 P4	131 40 33 39 19	212 222 - - -	7 24	15,65 13,22	.880-02 .290-02	- x	.0749 .0915	x xx
Nachpr., Erschöpfung, Konzentration, Prüf.art	P1 - P4 P1 P2 P3 P4	131 40 33 39 19	1121 2221 - - -	3 20	12,63 9,59	.950-03 .140-02	x x	.0813 .0857	xx xxx
Nachpr., Erschöpfung, Verlauf (i=schlechter)	P1 - P4 P1 P2 P3 P4	131 40 33 39 19	- - - - -						
Nachpr., Erschöpfung, Verlauf, Prüf.art (HD=2)	P1 - P4 P1 P2 P3 P4	131 40 33 39 19	1112 - - - -	16	5,08	.280-03	xx	.0809	xxx
Erschöpfung, Konzentration, Verlauf (i=schl., 3=bess)	P1 - P4 P1 P2 P3 P4	131 40 33 39 19	221 - - - -	27	16,00	.430-02	-	.0957	xx
Erschöpfung, Konzentration, Verlauf, Prüf.art	P1 - P4 P1 P2-P4	131 40 2211 2211 -	221 - - - -	24 7	11,60 1,75	.500-03 .160-02	x x	.1038 .1372	xxx xx

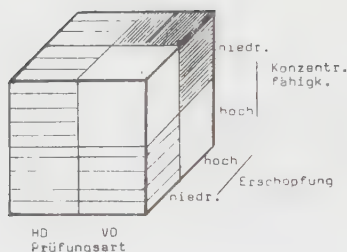


Abbildung 20: Erschöpfung und Konzentration nach der Prüfung, dreidimensionale Darstellung der KFA-Ergebnisse

Die beiden Antitypen bleiben erhalten und werden den VDN zugeordnet, wobei freilich der Prägnanzkoeffizient leicht fällt. Das Kollektiv der Erschöpften und Konzentrationsunfähigen wird bei wachsender Prägnanz zu 87 % den VDN zugeordnet. Wir haben versucht, dies in Abbildung 20 darzustellen.

Bei den Analysen der einzelnen Datenerhebungszeitpunkte gelingt es, bei P1 obige Ergebnisse eventuell erneut zu erhalten.

Die Variablen 'Verlauf der Prüfung' und 'Nachprüfung' (1 = zustimmen, 2 = ablehnen) sind mit den beiden anderen Variablen nur locker assoziiert. Dies ist daran erkennbar, daß die Zweier- und Dreier-KFAn entweder ergebnislos bleiben oder Typen berechnen, deren Prägnanz deutlich niedriger ist.

Die Anthrac-Klasse der HDn finden wir in Tabelle 34 nur zweimal. Es handelt sich um ein Kollektiv von

HDn, die einer Nachprüfung zustimmen würden, nicht erschöpft sind und angeben, die Prüfung sei schlechter als erwartet gelaufen.

Das wichtigste Ergebnis der Analogskalen stammt jedoch aus der Dreier-KFA mit den Variablen 'Erschöpfung', 'Konzentrationsfähigkeit' und 'Prüfungsart'. Das Kollektiv der Erschöpft-Konzentrationsunfähigen wird - bei leichten Prägnanzgewinnen - der Klasse der VDn zugeordnet, während das Kollektiv der Konzentrationsfähigen-Nichterschöpften durch Hinzufügen der Variablen 'Prüfungsart' aufgespalten wird und nunmehr die Signifikanzgrenze verfehlt.

4.3.2.1.6 Analyse der fremdskalierten "Aufregung"

Die Beisitzer waren gebeten worden, im Anschluß an die Prüfungen das Beisitzerblatt auszufüllen. Keiner der Beisitzer verweigerte dies.

Zur Validierung der durch Selbstrating erhobenen Belastungsmessungen wurden den Beisitzern zwei neunstufige Analogskalen zur Erfassung der 'Aufregung zu Beginn' und der 'Aufregung am Ende der Prüfung' mit den jeweiligen Polen 'extrem ruhig' (=1) und 'extrem aufgeregt' (=9) im Beisitzerblatt vorgelegt.

Obwohl es sich bei den Beisitzern um wissenschaftlich tätige Diplompsychologen handelte, konnten wir nicht davon ausgehen, daß wir es mit trainierten Beobachtern zu tun haben. Wir mußten also einen Weg finden, die Validität der Beobachtungen Untrainierter

zu erfassen. Wir baten jeweils zufällig anwesende Prüfungszuschauer, ebenfalls ein Beisitzerblatt auszufüllen. Sie taten dies nach Beendigung der Prüfung, während der Kandidat seine Zensur erfuhr.

Prüfungszuschauer sind in der Regel Psychologiestudenten, die demnächst die Prüfung, bei der sie hospitieren, ablegen werden. Kein Zuschauer wurde von uns ausgeschlossen und kein Zuschauer hospitierte in einer Prüfung, weil unsere Erhebung stattfand.

Die so erhaltenen Daten korrelierten wir, indem wir aus Beisitzern und Zuschauern je einer Prüfung Beobachterpaare bildeten.

Dabei zeigte sich, daß nur für die Skala 'Aufregung zu Beginn' eine signifikante Beobachterübereinstimmung errechnet werden konnte; sie liegt - je nach Berechnungsverfahren zwischen $r=.545$ und $r=.674$. Detailliertere Angaben sind aus Tabelle 48 und den dazugehörigen Erläuterungen im Text zu entnehmen. Zur Validierung der Selbststratingskalen werden wir demnach nur die Variable 'Aufregung zu Beginn' einsetzen.

4.3.2.2 Agglutination der Datensätze zur Belastungsmessung

In den Tabellen 35 und 36 sind die Ergebnisse der multivariaten Analysen der Emotionstests wiedergegeben.

Aus datenorganisationstechnischen Gründen wurden z. T. diejenigen beiden Versuchspersonen, für die nur dreifach wiederholte Messungen vorlagen, aus den Analysen ausgeschlossen (N = 39).

Tabelle 35: Ergebnisse der KFA_n der mediandichotomisierten Emotionstests für jeden Datenerhebungszeitpunkt einzeln; für die Läufe mit N = 41 liegt der Median des Datenerhebungszeitpunktes BL zugrunde, bei allen anderen Läufen der Gesamtmedian

Variablen in der KFA	n	signif. Typ	F	E	B	s.	Q	s.
BDI/STAI/HIS -BL	41	111 222	12 13	5,51 4,76	.620-o2 x .520-o3 xx		.1829 xx .2274 xxx	
BDI/STAI/HIS/HoS -BL	41	1111 2222	11 12	2,82 2,37	.750-o4 xx .2o0-o5 xxx		.2142 xxx .25o2 xxx	
BDI/STAI/HIS -BL	39	111 222	21 8	12,28 1,16	.310-o2 x .160-o4 xxx		.3263 xxx .18o8 xxx	
-P1	39	111 222	13 12	5,52 4,27	.2o0-o2 x .680-o3 xx		.2234 xx .2225 xxx	
-P2	39	111 222	1o 16	2,49 8,28	.130-o3 xx .410-o2 x		.2o58 xxx .2512 xx	
-P3	39	111	12	5,96	.110-o1 -		.1827 x	
-P4	19	-						
BDI/STAI/HIS/ Prüf.art -BL	41	2221	11	3,57	.350-o3 xx		.2o28 xxx	
BDI/STAI/HIS/HoS/ Prüf.art -BL	41	2221	1o	1,64	.4o0-o5 xxx		.1123 xxx	
BDI/STAI/HIS/ Prüf.art -BL	39	2221	7	o,8o	.130-o4 xxx		.1623 xxx	
-P1	39	-						
-P2	39	1112	5	o,76	.960-o3 x		.11o8 xxx	
-P3	39	-						
-P4	19	-						

Tabelle 36: Ergebnisse der KFA der mediandichotomisierten Emotionstests bei Verrechnung über alle Datenerhebungszeitpunkte; verwendet wurde der Gesamtmedian

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. Typ	F	E	S	s.	Q	s.
BDI/STAI/HIS	P1 - P4	140	111	41	16.52	.220-o7	xxx	.1983	xxx
			211	7	17.49	.250-o2	x	.0856	x
			112	8	17.00	.890-o2	=	.0731	x
			122	8	17.49	.650-o2	=	.0775	x
			222	42	18.52	.180-o6	xxx	.1933	xxx
BDI/STAI/HIS/Prüf.ert	P1 - P4	140	1111	26	11.21	.470-o4	xxx	.1148	xxx
			1112	15	5.31	.300-o3	xx	.0719	xxx
			2221	30	12.57	.610-o5	xxx	.1368	xxx

In der oberen Hälfte der Tabelle 35 sind die Ergebnisse der multivariaten Analysen für die einzelnen Datenerhebungszeitpunkte dargestellt. Im ersten Viertel liegt der Dichotomisierung der jeweilige Median des Datenerhebungszeitpunktes BL, im zweiten Viertel der jeweilige Gesamtmedian zugrunde. Übereinstimmend finden wir - vor allem in BL, P1 und P2 - Typen mit allgemein unterdurchschnittlichen Merkmalsausprägungen und Typen mit allgemein überdurchschnittlichen Merkmalsausprägungen. Inhaltlich entsprechen sie Kollektiven von allgemein 'Niedrig-Emotionalisierten' und allgemein 'Hoch-Emotionalisierten'.

Wiederholungsläufe unter Hinzufügung der Variablen Prüfungsart (VD/HD) brachten die Ergebnisse der unteren Hälfte der Tabelle 35. Danach rekrutieren sich die 'Hoch-Emotionalisierten' zum Zeitpunkt BL fast vollständig aus dem Kollektiv der VDN: das Kollektiv der 'Niedrig-Emotionalisierten' wird jetzt nicht

mehr als 'überzufällig' berechnet, hat sich demnach aufgespalten. Beim Zeitpunkt P2 fanden wir einen Typ niedrig-emotionalisierter HDn, der jedoch - bei deutlichen Pränanzverlusten - nur für 50 % des Gesamtkollektivs der Niedrig-Emotionalisierten bei diesem Datenerhebungszeitpunkt steht.

In Tabelle 36 sind die Ergebnisse von KFAn wiedergegeben, die bei Verrechnung über alle Prüfungsdatenerhebungszeitpunkte erhalten wurden. Wieder finden wir (bei fast unveränderter Pränanz) zwei signifikant überfrequentierte Typen: den Typ der Niedrig-Emotionalisierten und den Gegentyp der Hoch-Emotionalisierten. Zusätzlich finden wir in dieser Analyse jedoch drei Antitypen, d. h. Konfigurationen, deren Frequenzen signifikant unterbesetzt ausfallen. In Abbildung 20 haben wir versucht, die Ergebnisse der Tabelle 36 graphisch darzustellen.

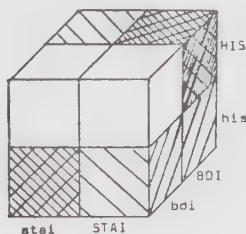


Abbildung 20: Dreidimensionale Darstellung der Ergebnisse der Tabelle 36; stark schraffiert bedeutet signifikant überbesetzt; die nicht sichtbare Zelle ist signifikant unterbesetzt

Die Hinzufügung der Variablen Prüfungsart ergibt zwei Typen für das Kollektiv der Niedrig-Emotionalisierten: niedrig-emotionalisierte VDN und niedrig-emotionalisierte HDN; ferner ergibt sich ein Typ hoch-emotionalisierter VDN; keiner der Antitypen 'überlebt' den Test gegen die Nullhypothese bei totaler Alpha-Adjustierung.

Tabelle 37: Ergebnisse der KFAn für die Differenzen der Datenerhebungszeitpunkte; die Stufe 2 der mit dem Wert 0.0 dichotomisierten Differenzen steht für gestiegene Scores, die Stufe 1 für gesunkene

Variablen in der KFA		N	signif. typ.	F	E	B	o.	Q	g.
BOI/STAI/HIS	BL-P1	41	222	19	11.39	.820-o2 -		.2615 x	
	P1-P2	39	111	11	3.60	.580-o3 xx		.2121 xxx	
			222	16	6.08	.120-o3 xxx		.3060 xxx	
	P2-P3	39	111	16	7.95	.250-o3 xx		.3289 xxx	
			222	9	2.49	.630-o3 xx		.1809 xxx	
	P3-P4	19	111	7	1.57	.540-o3 xx		.3115 xxx	
			222	9	3.32	.260-o2 x		.3621 xx	
BOI/STAI/HIS/ Prüf.art	BL-P1	41	-						
	P1-P2	39	1111	9	2.49	.640-o3 x		.1808 xxx	
			2221	12	4.21	.570-o3 xx		.2273 xxx	
	P2-P3	39	1111	13	5.51	.190-o2 x		.2271 xx	
			2221	4	0.74	.570-o2 -		.1784 xxx	
	P3-P4	19	1112	4	0.74	.580-o2 -		.2464 xx	
			2221	6	1.75				

Die Ergebnisse der KFAn für die Differenzwerte sind in Tabelle 37 dargestellt. Wir bildeten für jede Versuchsperson Differenzwerte, indem wir von dem erhaltenen Score den des jeweils vorhergehenden Datenerhebungszeitpunktes dann subtrahierten, wenn beide Scores erhoben worden waren. Wir zogen

also z. B. von HiS (P1) den Wert HiS (BL) ab; erhielten wir als Differenz eine positive Zahl, so war der HiS-Score gestiegen usw.; in Tabelle 37 liegt der Wert 0.0 der Dichotomisierung zugrunde.

Die starke Kovariation der Skalen zur Messung der Depressivität, Angst und Hilflosigkeit bei der Analyse der Testwerte fanden wir auch bei der Analyse der Testwertdifferenzen. Anders als bei der Analyse der Testwerte finden wir hier für die BL-P1-Daten nur einen (heuristischen) Typ auf dem 10-%-Signifikanzniveau, bei den anderen Differenzen prägnante Typen, die 69 % bzw. 84 % (P3 - P4) klassifizieren. Keine einzige der anderen möglichen Konfigurationsfrequenzen fällt in das Vertrauensintervall der in Tabelle 37 dargestellten signifikanten Typen.

Wir können daher aussagen, daß Veränderungen in einer Emotionsskala überzufällig häufig gleichsinnige Veränderungen in den anderen Emotionsskalen mit sich bringen.

Erneute KFAn unter Hinzunahme der Variablen Prüfungsart brachten nichts Neues, sondern die Ergebnisse sind analog den Ergebnissen der Testwertanalysen.

In Tabelle 38 sind die Ergebnisse der gemeinsamen Analyse der LIKERT-skalierten Emotionstests HiS, STAI und BDI und der Beschwerdelistenfaktorenscores dargestellt. Zur Dichotomisierung der Daten wurde der jeweilige Gesamtmedian verwendet.

**Tabelle 38: Ergebnisse der KFA mit den Emotions-
tests und den Faktorenscores der Be-
schwerdeliste**

Variablen in der KFA	Datenern. zeitpkt.	N	signif. - yp	F	E	B	s.	z	s.
HIS,STAI,BDI,BL1,BL2	1 - P4	140	11111 22212 22222 11112 11111 - -	18 12 13 7 7 - -	3,06 4,64 3,69 1,44 0,87	.350-00 .260-02 .940-04 .225-02 .430-04	xxx - xx - xxx	.3955 .0140 .0687 .1115 .1420	xxx xxx xxx xxx xxx
HIS,STAI,BDI,BL1,BL2, Prüf.art	1 - P4	140	11111 11112 22222 - 11111 - -	11 7 12 - 5	1,43 1,62 2,50 - 3,50	.720-03 .130-02 .940-05 - .710-03	x - xxx - x	.1115 .0140 .0687 - .1420	xxx xxx xxx - xxx

Bei der Analyse über alle Prüfungsdatenerhebungs-
zeitpunkte erhielten wir zwei hochsignifikante
Gegentypen: einen Typ mit ausschließlich unter-
durchschnittlichen Werten und einen Typ mit aus-
schließlich überdurchschnittlichen Werten. Auf
dem 10-%-Niveau - und damit als Hinweis zu werten -
findet sich ein Kollektiv, das in vier Variablen
überdurchschnittliche Werte, im ersten Beschwerde-
listenfaktor ('Angstsyndrom') jedoch unterdurch-
schnittliche Werte skaliert.

Bei Hinzufügung der Variable Prüfungsart (VD=1,
HD=2) zeigt sich, daß sich das Kollektiv mit aus-
schließlich unterdurchschnittlichen Skalierungen
sowohl aus den VDN als auch aus den HDN rekrutiert,
während das Kollektiv mit ausschließlich über-
durchschnittlichen Werten zu 92 % aus der Klasse

der VDN stammt. Ein Vergleich der Konfigurationsfrequenzen ergibt, daß VDN ebenso häufig unterdurchschnittliche wie überdurchschnittliche Werte skalieren, während die Konfigurationsfrequenz mit den Merkmalsausprägungen 'HD' und überdurchschnittliche Werte kleiner ist als die drei anderen.

Gegenüber der multivariaten Einzelanalyse der Emotionstests (vgl. Tabelle 36) zeigt sich allgemein ein Prägnanzverlust um mehr als 50 %.

Insgesamt scheinen die Typen eine mäßige, positive, lineare Assoziation zwischen den Emotionstests und den Beschwerdelistenfaktorenscores abzubilden.

In Tabelle 39 sind die Ergebnisse der gemeinsamen Analyse von Emotionstests und den beiden verbliebenen Analogskalen dargestellt.

Tabelle 39: Ergebnisse der KFAN mit den Emotionstests und den Analogskalen aus dem Fragebogen AN

Variablen im Test-KFAN	Unterschieds- zeitpunkt	n	signif. Typ	1	2	3	4	5	6
HIS, STAI, PDI, Erschöpfung, Konzentz.	P1 - P2	131	111111 222222	12 17	3.05 2.31	.140-04 .430-06	xxx xxx	.12-16 .12-16	xxx xxx
HIS, STAI, BDI, Erschöpfung, Konzentz., Frühp, art	P1 - P4	131	111111 111112 222221	13 7 14	3.05 1.16 2.93	.140-04 .160-03 .110-15	xxx x xxx	.07-11 .0-50 .14-11	xxx xxx xxx

Wir finden zwei jeweils hochsignifikante und prägnante Typen: ein Kollektiv mit unterdurchschnitt-

lichen Emotionen und ein Kollektiv mit überdurchschnittlichen Emotionen. Gegenüber den Einzelanalysen zeigt sich kein Prägnanzverlust um mehr als 50 %.

Die Hinzufügung der Variable Prüfungsart bringt eine Aufspaltung der Niedrig-Emotionalisierten in je ein Kollektiv VDN und HDN mit sich, während die Hochemotionalisierten zu 82 % ein Kollektiv von VDN darstellen.

In den Analysen über die einzelnen Datenerhebungszeitpunkte (deren Ergebnisse wir hier nicht dargestellt haben, weil sie redundant wären,) zeigte sich, daß 55 % des Gesamtkollektivs der Niedrig-Emotionalisierten und 61 % des Teilkollektivs der niedrig-emotionalisierten VDN sich aus dem Datenerhebungszeitpunkt P1 rekrutieren.

In Tabelle 40 sind die Ergebnisse der gemeinsamen Analyse der Beschwerdelistenfaktorenscores und der Analogskalen wiedergegeben.

Tabelle 40: Ergebnisse der KFAN mit den Faktorscores der Beschwerdeliste und den Analogskalen aus dem Fragebogen AN

Variablen der KFAN	Datenerh. zeitpkt.	N	signif. TVC	F	S	B	s.	D	s.
B1, B2, B3, B4, B5, B6, B7, B8, B9, B10, B11, B12, B13, B14, B15, B16, B17, B18, B19, B20, B21, B22, B23, B24, B25, B26, B27, B28, B29, B30, B31, B32, B33, B34, B35, B36, B37, B38, B39, B40, B41, B42, B43, B44, B45, B46, B47, B48, B49, B50, B51, B52, B53, B54, B55, B56, B57, B58, B59, B60, B61, B62, B63, B64, B65, B66, B67, B68, B69, B70, B71, B72, B73, B74, B75, B76, B77, B78, B79, B80, B81, B82, B83, B84, B85, B86, B87, B88, B89, B90, B91, B92, B93, B94, B95, B96, B97, B98, B99, B100	P1 = P2	31	1111 2222	21 14	11,30 5,76	.420-02 + .200-02 x		.0311 xxx .0859 xxx	
B1, B2, B3, B4, B5, B6, B7, B8, B9, B10, B11, B12, B13, B14, B15, B16, B17, B18, B19, B20, B21, B22, B23, B24, B25, B26, B27, B28, B29, B30, B31, B32, B33, B34, B35, B36, B37, B38, B39, B40, B41, B42, B43, B44, B45, B46, B47, B48, B49, B50, B51, B52, B53, B54, B55, B56, B57, B58, B59, B60, B61, B62, B63, B64, B65, B66, B67, B68, B69, B70, B71, B72, B73, B74, B75, B76, B77, B78, B79, B80, B81, B82, B83, B84, B85, B86, B87, B88, B89, B90, B91, B92, B93, B94, B95, B96, B97, B98, B99, B100	P1 = P2	13	2222	10	4,19	.310-03 xx		.0646 xxx	

Wir finden zwei gegenläufige Typen: ein Kollektiv mit überdurchschnittlichen Werten und ein Kollektiv mit unterdurchschnittlichen Werten, beide jeweils mäßig prägnant. Bei Hinzufügung der Variablen Prüfungsart (VD=1) verschwindet das Kollektiv der Unterdurchschnittlichen, das Kollektiv der Überdurchschnittlichen wird bei gleichbleibender Prägnanz der Merkmalsausprägung VD zugeordnet.

Gegenüber der Einzelanalyse der Analogskalen finden sich deutliche Prägnanzverluste, bei den Überdurchschnittlichen sogar um mehr als 50 %. Andererseits gelingt es, die Faktorenscores Typen zuzuordnen, was bei der Einzelanalyse nicht gelang.

In Tabelle 41 sind die Ergebnisse der gemeinsamen Analysen der Selbstrating-Verfahren und der fremdskalierten 'Aufregung zu Beginn', sowie die Ergebnisse der Wiederholungsläufe unter Hinzufügung der Variablen 'Prüfungsart' dargestellt.

Tabelle 41: Ergebnisse der KFA mit der fremd-beobachteten 'Aufregung zu Beginn' und den Selbstratingverfahren

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. Typ	F	E	B	s.	Q	s.
Aufregung zu Beginn, H15, STA1, BDI	P1 - P4	140	1111	22	8,32	.290-04	xxx	.1047	xxx
			1112	1	8,68	.130-02	=	.0590	xx
			1222	18	9,20	.490-02	-	.0878	xx
			2111	19	8,20	.560-03	xx	.0826	xxx
			2211	1	8,32	.190-02	=	.0560	x
			2222	24	9,07	.120-04	xxx	.1149	xxx
Aufregung zu Beginn, Brachöpf., Konzentration	P1 - P4	140	212	7	16,13	.360-02	=	.0802	x
Aufregung zu Beginn, BL1, BL2	P1 - P4	140	-						
Aufregung zu Beginn, H15, STA1, BDI, Prüf.ert	P1 - P4	140	11111	15	5,62	.540-03	x	.0303	xxx
			11121	0	5,87	.250-02	=	.0441	x
			22221	18	6,13	.440-04	xx	.0893	xxx
Aufregung zu Beginn, Brachöpf., Konzentration, Prüf.ert	P1 - P4	140	1211	3	11,66	.220-02	x	.0732	x
			2221	19	9,69	.370-02	-	.0774	xx
Aufregung zu Beginn, BL1, BL2, Prüf.ert	P1 - P4	140	-						

In der Agglutinations-KFA der drei Emotionstests mit der fremdskalierten Aufregung des Kandidaten zu Beginn der Prüfung erhielten wir zwei hochsignifikante und prägnante Typen, die Kollektive mit allgemein unterdurchschnittlichen bzw. allgemein überdurchschnittlichen Werten darstellen. Ferner erhielten wir zwei Antitypen, die uns bereits aus der multivariaten Analyse der Emotionstests bekannt sind (vgl. Tab. 36). Drittens fanden wir einen Hinweis auf ein Kollektiv mit unterdurch-

schnittlicher Aufregung und überdurchschnittlichen Emotionen. Viertens ließ sich ein Typ mit den Merkmalen überdurchschnittliche Aufregung und unterdurchschnittliche Emotionen sichern.

Bei Hinzufügen der Variablen Prüfungsart fallen nur noch drei Konfigurationsfrequenzen signifikant aus, ein Antityp lediglich auf dem 10-%-Niveau. Die beiden anderen Typen sind Kollektive von VDn, die fremd- und selbstskalierte überdurchschnittliche und fremd- und selbstskalierte unterdurchschnittliche Emotionen haben.

Dies deutet an, daß eine Assoziation zwischen selbst- und fremdskalierten Emotionen für das Kollektiv der VDn besteht, für das Kollektiv der HDn hingegen nicht.

In der Agglutinations-KFA mit der fremdskalierten Aufregung und den nach der Prüfung skalierten Variablen 'Erschöpfung' und 'Konzentrationsfähigkeit' sowie des Wiederholungslaufes unter Hinzufügen der Variablen 'Prüfungsart' registrieren wir drastische Prägnanzverluste. Deswegen müssen wir eine Assoziation ausschließen.

In keinem der Läufe über alle Datenerhebungszeitpunkte sowie für die einzelnen Datenerhebungszeitpunkte sowie bei Wiederholungsläufen unter Hinzufügen der Variablen 'Prüfungsart' fand sich bei der Agglutination der 'Aufregung zu Beginn' und den Faktorscores der Beschwerdeliste ein signifikanter Typ.

4.3.2.3 Zusammenfassung und Überblick

Die folgende überblicksartige Zusammenfassung gliedert sich in vier Abschnitte. Im ersten Abschnitt gehen wir auf die multivariaten Analysen der Emotionstests ein. Im zweiten Abschnitt greifen wir die zur Validierung der Emotionstests eingesetzten Datensätze und die Ergebnisse der Agglutinations-KFAn auf. Im dritten Abschnitt beziehen wir uns auf die Wiederholungsläufe unter Hinzufügen der Variablen Prüfungsart. Im vierten Abschnitt wiederholen wir die Hypothesen 6 und 8 und beziehen sie auf unsere Ergebnisse.

KFAn, mit denen wir die (formale) Reliabilität der Emotionstests HiS, STAI und BDI überprüften, erbrachten keinen signifikanten Typ. Hochsignifikante und prägnante Typen erhalten wir jedoch, wenn wir die 3 Tests multivariat verrechnen: bei der Analyse der mediandichotomisierten Rohwerte erhalten wir je einen Typ mit überdurchschnittlichen und je einen Typ mit unterdurchschnittlichen Merkmalsausprägungen, ängstliche Prüfungskandidaten sind also zugleich depressiv und hilflos, mutige zugleich 'euphorisch' und 'kompetent'. KFAn mit den Differenzwerten erbrachten außer für die Differenz der beiden ersten Datenerhebungszeitpunkte signifikante und prägnante Typen im Sinne gleichsinniger Änderungen aller Skalen: Kandidaten, die bei der nächsten Prüfung ängstlicher werden, werden zugleich depressiver und hilfloser und umgekehrt. Die fehlende (formale) Reliabilität einerseits und die prägnanten Typen aus den Analysen der

Rohwerte und der Differenzwerte andererseits zeigen, daß Prüfungsemotionen (gemessen mit dem STAI, dem BDI und der HiS) fluktuieren und nicht stabil sind, aber ein stets gleiches Muster ausbilden.

Das gleiche Ergebnis erhielten wir bei der Analyse der Beschwerdeliste: Die Faktorenstruktur ist reliabel, die Messungen der einzelnen Meßwertträger hingegen nicht.

In den Agglutinations-KFAn finden wir Übereinstimmend, daß sich die unterdurchschnittlichen Merkmalsausprägungen einerseits und die überdurchschnittlichen Merkmalsausprägungen andererseits jeweils zu Typen 'unterdurchschnittlicher Belastetheit' und 'überdurchschnittlicher Belastetheit' zusammenfinden. Dies ist zum Teil sehr überzeugend - wie bei der gemeinsamen Analyse der drei Emotionstests und den Analogskalen Erschöpfung und Konzentrationsfähigkeit -, z. T. weniger deutlich - wie bei der Agglutinations-KFA der drei Emotions-Tests und der fremdskalierten 'Aufregung zu Beginn', bei der jedoch die deutlich höheren Prägnanzkoeffizienten und das deutlich höhere Signifikanzniveau obige Zusammenfassung zulassen.

In diesem Sinne Übereinstimmend finden wir ebenfalls, daß bei Hinzufügen der Variablen Prüfungsart das Kollektiv der 'unterdurchschnittlich Belasteten' aufgespalten wird, während die 'überdurchschnittlich Belasteten' der Merkmalsausprägung VD zugeordnet werden.

Hypothese 6 lautete: "Die Indikatoren von Angst und Depression korrelieren hoch positiv und bilden gemeinsam die Symptomatik der Prüfungsbelastung". Aufgrund der überzeugenden Ergebnisse der multivariaten Analyse der drei Emotionstests HiS, STAI und BDI nehmen wir diese Hypothese an.

Hypothese 8 lautete: "Kandidaten in Zwischenprüfungen leiden unter stärkeren Prüfungsbelastungen als Kandidaten in Abschlußprüfungen". In formalem Sinne heißt dies, daß wir einen Typ mit der Merkmalsausprägung VD und überdurchschnittlichen Belastungswerten und einen zweiten Typ mit der Merkmalsausprägung HD und unterdurchschnittlichen Belastungswerten vorhersagen. Sowohl in der multivariaten Analyse wie in allen Agglutinations-KFAn, in die die drei Emotionstests eingingen, errechneten wir die vorhergesagten Typen als signifikant. Da wir sonst keine Ergebnisse erhielten, die diesem widersprechen, nehmen wir Hypothese 8 an. Posthoc können wir jedoch eine differenziertere Aussage machen: das Kollektiv der 'unterdurchschnittlich Belasteten' wird durch Hinzufügen der Variablen 'Prüfungsart' aufgespalten; dies bedeutet, daß HDn ihr 'Privileg' mit einer kleinen Gruppe VDn teilen müssen, während das überdurchschnittliche Kollektiv sich unzweifelhaft aus der Klasse der VDn rekrutiert.

4.3.3 Analyse der Strategien (Hypothesen 3, 5 und 9)

Hypothese 3 lautete: "Kandidaten in Zwischen- und in Abschlußprüfungen unterscheiden sich in konkret angebbaren Strategien und Verhaltensweisen." Als 5. Hypothese hatten wir formuliert: "Kandidaten in Zwischenprüfungen realisieren häufiger die Strategie der Gelernten Hilflosigkeit als Kandidaten in Abschlußprüfungen." Und in Hypothese 9 haben wir postuliert, daß es mindestens zwei multivariate Profile gibt, wovon das eine einer Kontrollstrategie, das andere der Strategie der Gelernten Hilflosigkeit entspricht und wovon das eine die Umkehrung des anderen ist.

In den ersten Abschnitten werden wir zunächst die einzelnen Datensätze (Meßinstrumente) untersuchen. Am Ende dieses Kapitels beziehen wir unsere Ergebnisse dann auf die o. a. Hypothesen.

4.3.3.1 Analyse der einzelnen Meßinstrumente

4.3.3.1.1 Das Testverhaltensdifferential (TVD)

Das Testverhaltensdifferential hatten wir als standardisiertes Verfahren unter zwei Instruktionsbedingungen eingesetzt:

1. zur Skalierung des vergangenen realisierten Prüfungsverhaltens,

2. zur Skalierung des zukünftigen Prüfungsverhaltens.

Die erste Fassung entspricht dabei der Originalversion.

Zunächst überprüfen wir mit den BL-Daten faktorenanalytisch die Dimensionierung. Wir extrahierten je drei Faktoren.

In Tabelle 42 sind die Ergebnisse der Faktorenanalysen wiedergegeben. Die unterstrichenen Pole sind dabei diejenigen, die bei der Scorebildung in Richtung erhöhter Anspannung, Variabilität und Unifizierung verrechnet werden.

Tabelle 42: Faktorenanalyse der TVD-Items für den Datenerhebungszeitpunkt BL

Dimension	Variable (Kurztext)	Instruktion: "werde"			"Habe"		
		F1	F2	F3	F1	F2	F3
Anspannung:	<u>besonnen</u> /zügig vorgehen		.06		-.15		
	<u>leichtfüßig/hartnäckig</u> vorgehen		-.72		-.77		
	rasch/ <u>grünlich</u> vorgehen		-.65		-.46		
	<u>anstrengen</u> /gelassen bleiben	-.74			.75		
	unbekümmert/ <u>sorgsam</u> vorgehen	.59	-.54		-.78		
Variabilität	<u>phantasievoll</u> /nüchtern denken		-.72			.61	
	in den Griff nehmen/ <u>auf mich wirken lassen</u>			.67	.27		
	systematisch/ <u>ohne Systematik</u> vorgehen	.40	.70		.79		
	<u>variabel</u> /stetig vorgehen	-.77			-.65	.46	
	<u>Einfälle aufsteigen lassen</u> /unterdrücken	-.66				.73	
Unifizierung	<u>ganzheitlich auffassen</u> /zergliedern		.64		.70		
	Zusammenhänge auflösen/ <u>herstellen</u>	.45	-.48			.86	
	<u>Ganzen</u> /Einzelheiten betonen	-.55	.47			-.85	
	<u>Kern der Sache</u> /Einzelheiten beachten		.69		.61		
	<u>verknüpfen</u> /trennen	-.73			.54	-.59	
Aufgeklärte Varianz (in %)		17.0	20.4	16.4	25.5	17.7	15.0

Zunächst einmal ist festzuhalten, daß die drei Dimensionen faktorenanalytisch keineswegs überzeugend abgebildet werden, obwohl nach dem Scree-Test jeweils drei Faktoren zu extrahieren sind (Eigenwertverlauf hier nicht dargestellt).

Auf dem stärksten Faktor der 'Werde'-Daten mit 20,4 % aufgeklärter Varianz laden die Variablen

- besonnen, hartnäckig, gründlich, sorgsam, systematisch und stetig vorgehen, nüchtern denken.

Faktor: Stetigkeit und Konzentration.

Auf dem zweitstärksten Faktor mit 17,0 % aufgeklärter Varianz laden die Variablen

- sich anstrengen, sorgsam vorgehen, Einfälle aufsteigen lassen, Zusammenhänge herstellen, Ganzes betonen, Dinge verknüpfen.

Faktor: Synthetische Kreativität.

Auf dem drittstärksten Faktor mit 16,4 % aufgeklärter Varianz laden die Variablen

- Dinge auf mich wirken lassen, ohne Systematik vorgehen, Dinge zergliedern, Zusammenhänge auflösen, Einzelheiten betonen, beachten und trennen.

Faktor: Passiv-analytisches Vorgehen.

Auf dem stärksten Faktor der 'Habe'-Daten mit 25,5 % aufgeklärter Varianz laden die Items:

- zügig, hartnäckig, gründlich, sorgsam, systematisch und stetig vorgehen, sich anstrengen, die Dinge in den Griff nehmen.

Faktor: Stetigkeit und Konzentration.

Auf dem zweitstärksten Faktor laden die Items

- stetig vorgehen, Einfälle unterdrücken, nüchtern denken, Dinge zergliedern, Einzelheiten beachten und trennen.

Faktor: Analytisches Vorgehen.

Auf dem drittstärksten Faktor mit 15,0 % aufgeklärter Varianz laden die Items

- Zusammenhänge herstellen, Ganzes betonen, Dinge verknüpfen.

Faktor: Synthetisches Vorgehen.

Die hier nicht dargestellten Faktorenanalysen der Datenerhebungszeitpunkte P1 bis P4 ergeben ähnliche Ergebnisse: Die inhaltliche Interpretation der Faktoren läßt sich unschwer auf die Dimensionen der Anspannung, der Unifizierung/Diversifizierung und der Variabilität/Stetigkeit beziehen. Die Ladungsmatrizen hingegen zeigen so wenig Ähnlichkeiten, daß wir auf Ähnlichkeitsrotationen verzichten. Das Nahziel - kritische Analyse eines nicht faktorenanalytisch konstruierten Tests - ist somit erreicht: Methodisch gesehen mißlang es, über Faktorenanalysen die Dimensionen des Test zu gewinnen, und die so erhaltenen Lösungen sind auch nicht reliabel.

Dennoch rechneten wir im folgenden mit den vorschriftsmäßig errechneten Summenscores weiter (zur Verrechnung vgl. Programme im Anhang).

Dies mag auf den ersten Blick befremden: Tatsächlich haben wir jedoch (vgl. Kapitel 3.3.2) von vornherein adimensionale Strukturen vermutet, die evtl. nicht dem Modell der Faktorenanalyse - lineare und additive Beziehungen und Fehlen von Assoziationen höherer Ordnung - entsprechen. Es kann z. B. ein Fall vorliegen, in dem ein Teil der Kandidaten Strategien realisiert und intendiert, die mit dem TVD meßbar sind, ein anderer Teil aber nicht; die resultierenden Faktorenanalysen wären dann vermutlich ebenso 'unbefriedigend' ähnlich wie die oben dargestellten.

Im folgenden gehen wir daher (zunächst) davon aus, das TVD würde die Dimensionen Anspannung, Variabilität und Unifizierung messen.

In Tabelle 43 sind die erhaltenen deskriptiven Statistiken sowie die routinemäßig gerechneten t-tests zwischen VDn und HDn dargestellt.

Tabelle 43: Mittelwerte und Standardabweichungen der TVD-Testscores

Variable	Datenerh. zeitpkt.	M	s	M(VD)	M(HD)	p(t-test)
Anspannung(W)	BL	0.96	0.81	0.90	1.10	xx
	P1	0.30	0.88	0.65	1.17	
	P2	0.71	0.94	0.62	0.90	
	P3	0.68	1.00	0.41	1.30	
	P4	0.90	0.95	0.98	0.80	
Anspannung(H)	BL	0.42	0.86	0.32	0.65	-
	P1	0.39	0.92	0.38	0.40	
	P2	0.47	1.04	0.29	0.88	
	P3	0.23	0.96	0.05	0.62	
	P4	0.52	1.03	0.27	1.04	
Variabilität(W)	BL	-0.07	0.92	-0.01	-0.22	x x
	P1	-0.24	0.70	-0.13	-0.52	
	P2	-0.45	0.87	-0.26	-0.98	
	P3	-0.24	0.92	-0.04	-0.70	
	P4	-0.39	0.85	-0.36	-0.42	
Variabilität(H)	BL	0.10	1.03	0.41	-0.70	x - x x
	P1	-0.25	0.71	-0.12	-0.57	
	P2	-0.01	0.94	0.17	-0.43	
	P3	0.02	0.94	0.26	-0.52	
	P4	0.13	0.85	0.11	-0.02	
Unifizierung(W)	B1	1.14	0.81	1.21	0.99	
	P1	1.25	0.91	1.26	1.23	
	P2	1.52	0.95	1.65	1.22	
	P3	1.11	0.88	1.19	1.12	
	P4	1.25	0.94	1.40	1.00	
Unifizierung(H)	BL	0.43	1.09	0.39	0.50	-
	P1	0.83	0.94	0.68	1.22	
	P2	0.92	1.09	0.87	1.02	
	P3	0.68	1.12	0.77	1.12	
	P4	0.82	1.05	0.80	1.09	

Insgesamt ergibt sich, daß die Versuchspersonen ihr geplantes und realisiertes Verhalten auf der Dimension Anspannung eher in Richtung stärkerer Anspannung skalieren, auf der Dimension Variabilität/Stetigkeit eher in Richtung Stetigkeit und auf der Dimension Unifizierung/Diversifizierung eher in Richtung eines globalen, synthetischen Vorgehens.

Auffallend ist, daß das intendierte Verhalten jeweils stärker in die o. a. Richtung geht als das tatsächlich realisierte, obwohl die Werde- mit der Habe-Version relativ gut miteinander korrelieren.

In Tabelle 44 haben wir die t-tests für abhängige Stichproben und die dazugehörigen Korrelationen aufgeführt.

Tabelle 44: Korrelationen und Mittelwertsunterschiede zwischen den beiden Instruktionsbedingungen des TVD

	SL	L1	L2	W1	W2
A (W) v. A (H)	.70	.65	.70	.65	.70
U (W) v. U (H)	.65	.60	.65	.60	.65
U (W) v. U (H)	.60	.55	.60	.55	.60

Die Differenz der Werde-Habe-Werte betrachten wir als Maß für die Dissonanz zwischen Wunsch und Wirklichkeit.

In Tabelle 45 sind die signifikanten Korrelationen zwischen den 6 TVD-Werten und 3 Differenzwerten für N = 140 dargestellt.

Tabelle 45: Korrelationen der TVD-Testscores
über alle Prüfungs-Datenerhebungs-
zeitpunkte (1 bis P4)

Variablen des TVD								
A(W-H)	V(W-H)	U(W-H)	A(W)	V(W)	U(W)	A(H)	V(H)	U(H)
A(W-H)		.24	.47			-.51		
	V(W-H)			.50			-.51	
		U(W-H)			.41			-.63
			A(W)	-.44	.17	.52	-.44	
				V(W)		-.27	.49	-.19
					U(W)			.46
						A(H)	-.30	.18
							V(H)	

Erwartungsgemäß korrelieren die Differenzen mit ihren Minuenden und Subtrahenden. Übereinstimmend zeigt sich eine schwache positive Korrelation zwischen allen Werten der Dimension Anspannung und Unifizierung sowie - läßt man die Differenzwerte unberücksichtigt - eine negative Korrelation zwischen Anspannung und Variabilität.

Mit Konfigurationsfrequenzanalysen analysierten wir die multivariaten Zusammenhänge. Zugrunde legten wir jeweils die Gesamtmediane.

In Tabelle 46 sind die für die einzelnen Datenerhebungszeitpunkte gerechneten KFAn aufgelistet.

Tabelle 46: Ergebnisse der KFAn mit den mediandichotomisierten TVD-Scores für die einzelnen Datenerhebungszeitpunkte

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. Typ	F	E	B	s.	7	s.
A(w), v(w), u(w)	BL P1 P2 P3 P4	41 41 41 39 19	- - - 121 212	15 8	5.90 3.16	.180-01 .120-01	x x	.24 xx .171 x	
A(w), v(w), u(w), Prüf.art	BL P1 P2 P3 P4	41 41 41 39 19	- - - 1211 -	13	4.77	.503-03	xx	.2403 xxx	
In keiner der Dreier-KFAn mit den Variablen A(H), V(H), U(H) und in keiner der Vierer-KFAn mit den Variablen A(H), V(H), U(H), Prüf.art (V0=1, H0=2) ergab sich ein signifikanter Typ.									
A(w), v(w), u(w), A(H), V(H), U(H)	BL P1 P2 P3 P4	41 41 41 39 19	121121 - 121212 121121 -	6 6 8	1.35 0.71 1.12	.010-01 .120-01 .120-01	x x xxx	.237 xxx .134 xxx .1 xxx	
s.o., Prüf.art	BL P1 P2 P3 P4	41 41 41 39 19	- - - 121121 -	7	0.76	.100-04	xx	.1647 xxx	
A(w), v(w), u(w), P1, P2, P3, P4		19 19 19 19 19	- - - - -						
Alle 6 o.a. Fünfer-KFAn wurden als Sechser-KFAn unter Hinzufügung der Variable Prüfungsart wiederholt. Es ergab sich in keinem Lauf ein signifikanter Typ									
A(w)/(BL, P1, P2, P3)		39	1111	9	1.00	.200-04	xxx	.1975 xxx	
v(w)/		39	1111	10	1.17	.100-02	xxx	.2184 xxx	
j(w)/		39	2222	12	3.65	.160-01	xx	.2402 xxx	
u(w)/		39	-						
A(H)/		39	-						
V(H)/		39	1222	6	1.14	.120-01	x	.143 xxx	
U(H)/		39	-						
A(w)/(BL, P1, P2, P3, Prüf.art)		39	11111	8	1.25	.280-04	xxx	.1738 xxx	
v(w)/		39	11111	5	0.98	.280-02	-	.1100 xxx	
j(w)/		39	11112	5	0.96	.730-04	xxx	.1100 xxx	
u(w)/		39	12121	12	2.17	.400-05	xxx	.1975 xxx	
A(H)/		39	-						
V(H)/		39	22221	6	0.94	.330-03	xx	.1749 xxx	
U(H)/		39	-						

Bei der multivariaten Analyse der jeweils drei Dimensionen erhielten wir lediglich für die intendierten Verhaltensweisen beim Datenerhebungszeitpunkt P3 signifikante Ergebnisse. Rund 59 % aller Versuchspersonen werden einer von zwei Typen zugeordnet: den zukünftig mäßig Angespannten, Variablen, Differenzierend-Analytischen und den zukünftig Angespannten, Stetigen, global-synthetisch Vorgehenden. Bei Hinzufügung der Variablen Prüfungsart zeigt sich, daß erstere Gruppe ein Kollektiv von VDn darstellt.

Bei der multivariaten Analyse sowohl der realisierten als auch der intendierten Strategien zeigt sich, daß realisierte wie intendierte Strategien jeweils die gleiche Struktur haben und sich in entsprechenden Typen niederschlagen. Beim Datenerhebungszeitpunkt P3 sind diejenigen, die mäßig angespannt, variabel und diversifizierend vorgegangen sind und zukünftig auch vorgehen werden, wiederum ein Kollektiv von VDn.

Bei der Analyse der formalen Reliabilität zeigt sich, daß für die Analyse mit $N = 19$ über die Datenerhebungszeitpunkte BL, P1, P2, P3 und P4 die Stichprobe zu klein ist, um eine Konfigurationsfrequenz den Schritt über die alpha-adjustierte Signifikanzhürde schaffen zu lassen.

Bei der Analyse mit $N = 39$ über die Datenerhebungszeitpunkte BL, P2, P2 und P3 finden wir für die Dimension Unifizierung/Diversifizierung keinen Typ.

Bei der Analyse der Anspannungsdimension finden wir

einen Typ, der zukünftig mäßig angespannt vorgehen will; dieser Typ ist der Merkmalsausprägung VD der Variablen 'Prüfungsart' zuzuordnen.

Bei der Analyse der intendierten Dimension Variabilität-Stetigkeit lassen sich 59 % aller Versuchspersonen einem von zwei Typen zuordnen, den 'Stetigen' und den 'Variablen'. Bei Hinzufügung der Variable Prüfungsart zeigt sich, daß die 'Variablen' ein Kollektiv von 'variablen VDn' bilden.

Insgesamt ist die Dimension Variabilität reliabler als alle anderen Dimensionen; für die (formale) Reliabilität der Dimension Unifizierung fehlt jeder Anhaltspunkt. Ferner sind die intendierten Prüfungsstrategien reliabler als die realisierten.

Bei der Analyse der Differenzwerte - sowohl der zwischen den Datenerhebungszeitpunkten als auch der zwischen den beiden Instruktionen - ergab sich kein signifikanter Typ. In Tabelle 47 haben wir die multivariate Analyse der Testscores über alle Prüfungs-Datenerhebungszeitpunkte wiederholt.

Tabelle 47: Ergebnisse der KFA der mediandichotomisierten TVD-Scores über die Datenerhebungszeitpunkte P1 bis P4

Variablen in der KFA	N	signif. TVD	F	$\bar{z}$	A	z	Q	%
$A(U), V(H), U(H)$	140	121 212	13 25	20,59 14,60	.395-02 x .510-02 x		.1019 .0819	xx xx
$A(H), V(U), U(H)$	140	-						
$A(U), V(U), U(U)$ $A(H), V(H), U(H)$	140	121121 212212	17 14	2,43 1,88	.480-09 xxx .810-08 xxx		.1059 .0878	xxx xxx
$A(J-H), V(U-H), U(U-H)$	140	-						
$A(U), V(U), U(J), \text{Prüf.art}$	140	1211	27	13,97	.640-03 x		.1034	xxx
$A(H), V(H), U(H), \text{Prüf.art}$	140	2122	13	5,78	.550-02 -		.0538	xx
$A(U), V(U), U(U), \text{Prüf.art}$	140	1211211 2122121 2122122	14 9 5	1,65 1,27 0,60	.160-08 xxx .640-05 xxx .380-03 x		.0893 .0557 .0315	xxx xxx xxx
$A(U-H), V(U-H), U(U-H), \text{Prüf.art}$	140	-						

Es bestätigt sich, daß die intendierten Vorgehensweisen überzufällig strukturiert sind. Bei der gemeinsamen Analyse der intendierten und realisierten Vorgehensweisen zeigt sich, daß beide Datensätze die gleiche multivariate Struktur besitzen; obwohl die absolute Zahl der Merkmalsträger sich fast halbiert, ist diese Aussage zulässig, weil die Prägnanz gleich bleibt.

Bei Hinzufügung der Variablen 'Prüfungsart' zeigt sich, daß die zukünftig locker, variabel und analytisch Vorgehenden ein Kollektiv von VDN bilden. Dies gilt auch bei der multivariaten Verrechnung der intendierten und realisierten Strategien. Der Gegentyp, das Posthoc-Kollektiv der Angespannt-

stetig-synthetisch-Vorgehenden, wird hingegen aufgespalten.

Zusammenfassend ergibt sich auf der Basis der Ergebnisse der TVD-Analysen ein Bild, nach dem es in der untersuchten Stichprobe zwei Kollektive mit relativ reliablen multivariaten Strategien gibt, das Kollektiv der mäßig Angespannten, variabel und diversifizierend Vorgehenden und das Kollektiv der angespannt, stetig und unifizierend Vorgehenden. Diese Kollektive neigen dazu, ihre Strategien beizubehalten. Ein Teil der Versuchspersonen jedoch wechselt zwischen den Dimensionen oder hat gar keine Strategien, die mit dem TVD meßbar wären (so daß sich Zufallsdaten ergeben). Faktorenanalysen über wiederholt gemessene Daten können unter diesen Umständen keine stabile Struktur erreichen. Die (formale) Reliabilität der Einzelskalen deutet jedoch auf ein Kollektiv konstant Skalierender, und die Analyse über $N = 140$ (wobei 'unabhängige' und 'abhängige' Messungen in einem Rutsch verrechnet wurden,) ergibt dann durch mehrfache Verrechnung der Daten derselben Versuchsperson einen entsprechenden Typ.

4.3.3.1.2 Exkurs: Analyse der Daten, die von Fremdbeobachtern stammen - das Beisitzerblatt

Neben Items zur Validierung der Zensur und zur Validierung der durch Selbststrating erhobenen Belastungsmessungen enthielt das Beisitzerblatt

Items zur Beobachtung der Prüfungssteuerung durch Prüfer und Kandidaten (bezüglich des Kandidaten wollen wir von jetzt an den Begriff der 'Taktik' verwenden).

Auf die Problematik des Einsatzes untrainierter Beobachter sind wir bereits in Kapitel 4.3.2.1.6 eingegangen. Die von Beisitzern und zufälligen Prüfungszuschauern erhaltenen Daten korrelierten wir, indem wir aus Beisitzern und Zuschauern je einer Prüfung Beobachterpaare bildeten. In Tabelle 48 sind die Produkt-Moment-Korrelationen und Spearman-Rang-Korrelationen wiedergegeben. Die Daten der linken Tabellenhälfte beziehen sich auf NPAAR=31 Paare, d. h. jeder Zuschauer und jeder Beisitzer einer Prüfung wurde mit jedem korreliert, die rechte Hälfte gibt die Daten von 11 Beisitzern und 11 - ggf. per Los ermittelten - Zuschauern (in 11 Prüfungen waren also Zuschauer anwesend).

Tabelle 48: Interrater-Reliabilität der Fremdbeobachteritems

Item	Text	r(NF=31)	p	R(NP=31)	p	r(NP=11)	p	R(NP=11)	p
1.	Ist der Kandidat beim Antworten ausgeglichen durch:								
a)	Verwendung von Themen, die inhaltlich zur Prüfung gehören, jedoch nicht gefragt wurden	.022		.105		-.259		-.211	
b)	Reden über Gefühle	-.098		-.045		-.130		-.148	
2.	Wie häufig hat der Kandidat Fragen explizit zurückgewiesen oder spontan mit "weiß ich nicht" beantwortet?	.172		.308	x	.584	x	.647	x
3.	..Wie häufig hat der Prüfer den Kandidaten unterbrochen?	.462	xx	.485	xx	.670	x	.808	xxx
4.	Konnten Sie beim Prüfer während der Prüfung Gefühle beobachten?(Summe)	-.142		-.142		.418		.418	
6.	Bitte schätzen Sie jetzt noch ab, wie aufgeregt der Kandidat war:								
a)	zu Beginn	.545	xxx	.562	xxx	.674	xx	.551	x
b)	am Ende der Prüfung (Differenz a) - b)	.299		.270		.186		.171	
		-.065		-.064		-.101		-.174	

Es zeigt sich, daß allein für die Häufigkeit, mit der ein Prüfer den Kandidaten unterbrochen hat (unabhängig davon, warum er dies tat), eine befriedigende Beobachterübereinstimmung registriert werden kann.

Das Item der zurückgewiesenen Fragen erreicht in einem der Kennwerte kein akzeptables Signifikanzniveau, wohl aber in den 3 anderen. Wir wollen es deshalb als mäßig reliabel betrachten und beibehalten.

In den weiteren Analysen werden wir uns nur noch mit den reliablen Items beschäftigen.

Berechnungen von Unterschieden zwischen Vor- und Hauptdiplomanden, zwischen der Einschätzung verschiedener Beisitzer und zwischen verschiedenen Prüfern hinsichtlich der reliablen Items ergab ein signifikantes Ergebnis: einer der Prüfer unterbrach signifikant häufiger als alle anderen Prüfer ($p < .01$).

4.3.3.1.3 Analyse der Strategiensammlung (SS)

Neben den Testverhaltensdifferentialen setzten wir die 34 Items der Strategiensammlung (vgl. Fragebogen F3 im Anhang) zur Messung prospektiver Strategien ein.

Wegen der von uns vermuteten adimensionalen Struktur strebten wir eine konfigurationsfrequenzanalytische Verrechnung an. Auswertungspraktisch folgt daraus der Zwang zur Reduktion der Information auf wenige Kennwerte für Strategiedimensionen bzw. Strategiehierarchiee.

4.3.3.1.3.1 Ergebnisse von Faktorenanalysen des Typs R

Wir rechneten zunächst Faktorenanalysen des Typs R (vgl. HERRMANN, 1969, S. 158 ff.) getrennt für jeden einzelnen Datenerhebungszeitpunkt.

Die übereinanderprojizierten Eigenwertverläufe (vgl. Abbildung 21) zeigen, daß nach dem Scree-Test insgesamt je drei Faktoren zu extrahieren sind.

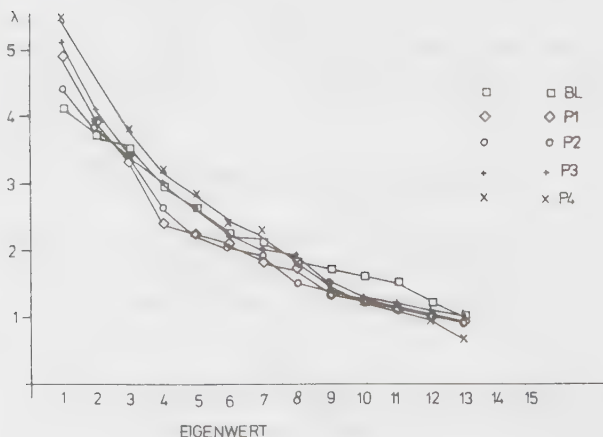


Abbildung 21: Übereinanderprojizierte Eigenwertverläufe aus Hauptkomponentenanalysen der SS-Items jedes Datenerhebungszeitpunktes

Tabelle 49 gibt die Ergebnisse der fünf Faktorenanalysen (BL, P1, P2, P3, P4) wieder. Aufgeführt sind nur Ladungen von $a > |.40|$.

Tabelle 49: Varimaxrotierte Hauptkomponentenanalysen der SS-Items mit paarweisem Ausschuß von Missing Values separat für jeden Datenerhebungszeitpunkt

Datenerh.	Baseline			P1			P2			P3			P4		
Item	F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3
1			.49			-.50	-.60			.44			-.50		
2	.58				.64			.62		.71				.65	
3	.60				.77			.75		.78			-.65		
4		.41			-.41		-.49				.43		.55		
5					.85			.59		.66			-.61		
6		-.45			-.66	.42					.76				
7			.45					.50		.41					
8			.44					-.53		.69			-.57		
9				-.48			.48						.49		
10			-.56				.51								
11		.64		.43				.49			.64		-.60		
12					.46		.57				.76				
13					-.41		.49	-.46			.56		.60		
14	.44				-.43			.49	.58			-.47			.44
15		.54	-.47	.41			.51						.86		
16							.56								
17		.53				.66				.47			.76		
18		.57				.59							.62		
19				-.73				-.62		.67				-.61	
20	.44							-.65		.49					
21					-.41								.64		
22		.62		-.70			-.60	.42		.56			.51		
23		.44								.53			.65		
24				.51			.52			-.74			.41	.54	
25				.45										-.66	
26	.56				.54		-.46						-.51		
27	.55			-.43				-.45					.46	.43	
28	.57			-.70			.52						.56		
29				-.70			-.43			.63			.44		
30			-.49	.47			.59			-.45					
31		-.62		.51			.55			-.53	-.47			.40	
32		.65			.59					.61			-.72		
33	.76				.52		.70				-.80		.55		
34		.55		.67			.64			-.68				.76	

Auf den ersten Blick erkennbar sind die fehlenden gemeinsamen Komponenten der fünf Matrizen. Nur zehn

Items erreichen substantielle Ladungen in allen Analysen; keine der 45 möglichen Zweierkombinationen dieser Items findet sich in allen fünf Analysen auf demselben Faktor!

Sieht man von den formalen Mängeln der Faktorenanalysen (34 Items bei 41/41/41/39/19 Versuchspersonen mit paarweisem Ausschluß von Missing Values) ab, geht man weiter davon aus, daß wir keine Zufallszahlen analysiert haben (Scree-Test!), so zeigen die offensichtlich unterschiedlichen Faktorenmatrizen, daß die Faktorenanalyse als Methode der Wahl in diesem Fall nicht angemessen ist.

Ähnlich diskrepante Lösungen erhielten wir übrigens, als wir in den Datensätzen die Missing Values durch Mittelwerte ersetzten. Wir haben diese Ergebnisse deshalb hier nicht dargestellt.

Wir müssen deshalb den auf S. 195 ff. beschriebenen Weg der Selektion geeigneter Items beschreiten.

4.3.3.1.3.2 Selektion von 'Kernstrategieitems'

Auf die Darstellung der Zwischenergebnisse dieses rechentechnisch sehr aufwendigen Weges verzichten wir, um die Übersichtlichkeit zu erhalten.

In einem ersten Schritt zur Identifikation der Kernstrategien berechneten wir Q-Analysen. Wir rechneten für jeden Datenerhebungszeitpunkt getrennt mit allen

41/41/41/39/19 Versuchspersonen und mit allen 34 Items (Missing Value=0) Hauptkomponentenlösungen mit anschließender Varimaxrotation. Die Zahl der zu extrahierenden Faktoren bestimmten wir mit dem Scree-Test. Bei Q-Analysen werden keine Items, sondern Versuchspersonen korreliert. Eine Korrelation zwischen Versuchspersonen von $r=1.0$ bedeutet, daß die beiden Versuchspersonen ein identisches Antwortprofil haben. Ein Q-Faktor steht dann für eine abstrahierte 'typische' Versuchsperson bzw. für ein 'typisches' Profil in der untersuchten Stichprobe.

Die absolute Höhe der Ladung auf den Personfaktoren nahmen wir als Maß der Ähnlichkeit zwischen Person und 'Profil', i. e. je höher die Ladung, desto 'typischer' die Versuchsperson.

Um Mißverständnissen vorzubeugen, möchten wir betonen, daß der 'Typ'-Begriff in der Q-Analyse nichts mit dem Typ-Begriff in der KFA zu tun hat. 'Typ' und Typ sollten nicht verwechselt werden.

In einem zweiten Schritt selegierten wir aus dem Itempool 18 Items, die entweder in allen fünf R-Analysen substantielle Ladungen ($a > |.40|$) oder in vier Analysen substantielle und in mindestens einer Analyse eine Ladung $a > |.70|$ erreichten.

Auf der Basis der Ergebnisse der Q-Analysen und der Selektion von 18 Markier-Items aus den R-Analysen rechneten wir weiter.

Jede Versuchsperson ordneten wir (innerhalb jedes Datenerhebungszeitpunktes) entsprechend ihrer

höchsten Ladung in den Q-Faktoren einer Gruppe zu. Mit den Gruppen als Faktor und mit den 18 Markier-Items als abhängige Variablen rechneten wir für jedes Item einfaktorielle Varianzanalysen. Dies ergab für 5 Datenerhebungszeitpunkte und 18 Variablen $5 \times 18 = 90$ Varianzanalysen!

Ein signifikanter F-Wert besagt, daß sich die durch die Q-Analysen gefundenen homogenen Gruppen in diesem Item unterscheiden. Damit ist dieses Item u. E. besonders geeignet, Typen und Gegentypen im Sinne der KFA abzubilden. Fielen die F-Tests in mindestens 2 Datenerhebungszeitpunkten signifikant aus, dann vergaben wir einen Gewichtungspunkt ("F-Test" in Tabelle 50).

Weiterhin berechneten wir auf jeden Q-Faktor eine schrittweise, multiple Regression mit den 34 Variablen des Itempools, wobei wir aus rechentechnischen Gründen die Missing Values durch den jeweiligen Variablenmittelwert ersetzten. Jede Variable, die in die Regressionsanalyse aufgenommen wurde, bevor die Änderung des Quadrats der multiplen Korrelation unter $\Delta r = .01$ sank, wurde in eine Korrelationsmatrix aufgenommen, die aus den Variablen als Zeilen, den Q-Faktoren als Spalten und den einfachen Korrelationen als Werten bestand.

Auf der Basis dieser Korrelationsmatrix vergaben wir einen Gewichtungspunkt, wenn ein Item in allen fünf Datenerhebungszeitpunkten mit mindestens einem Q-Faktor signifikant korrelierte (r_Q in Tabelle 50), und einen weiteren Gewichtungspunkt, wenn eine

Variable bei mehr als einem Datenerhebungszeitpunkt mit mindestens zwei Faktoren in umgekehrter Richtung signifikant korrelierte (r_Q +/- in Tabelle 50).

In Tabelle 50 haben wir die aus den R-Analysen selegierten Markieritems und die vergebenen Gewichtungspunkte dargestellt.

Tabelle 50: Gewichtungspunkte für die Markier-
 iters aus den SS-Analysen

Item	Text	F-Test	r _Q	r _Q (+/-)	Gewichtung
1	"Wenn ich anderer Meinung bin, dann werde ich auch in der Prüfung einem Professor widersprechen."	x	x	x	3
2	"Ich werde betont selbstbewußt auftreten."				0
3	"Ich werde mich optimistisch zeigen."				0
4	"Ich werde möglichst wenig sagen."	x			1
5	"Ich werde mich mutig zeigen."	x	x	x	3
11	"Ich muß alle Fragen sehr genau und gewissenhaft beantworten."				0
14	"Das Beste ist, wenn ich kurz und bündig auf die Fragen des Prüfers antworte."	x	x	x	3
15	"Wenn ich in einer Prüfung viel rede, erzähle ich auch viel Falsches."		x		1
17	"Wenn ich bei einer Frage die Antwort nicht weiß, dann heißt das, daß ich bei der Vorbereitung nicht aufgepaßt habe."		x	x	2
19	"In einem gewissen Rahmen kann ich auf den Prüfungsverlauf Einfluß nehmen."			x	1
22	"Ich gehe davon aus, daß mir in der Prüfung auch auf schwere Fragen eine Antwort einfallen wird."	x	x	x	3
24	"Prüfer stellen nur scheinbar leichte Fragen, in Wirklichkeit wollen sie mehr wissen."	x	x	x	3
28	"Es kommt darauf an, daß ich die Zusammenhänge und Unterschiede anschaulich darstelle."	x			1
29	"Es ist für die Zensur unerheblich, wenn ich ein paar Details nicht weiß."			x	1
31	"Es ist besser, wenn ich nicht allzu kritisch auftrete."	x	x	x	3
32	"Wenn ich an einer Theorie etwas auszusetzen habe, werde ich dies sehr deutlich zum Ausdruck bringen."			x	1
33	"Es ist wichtig, daß ich spontan antworte."			x	1
34	"Wenn ich bei mehr als 3 Fragen überlegen muß, macht das einen schlechten Eindruck."	x	x	x	3

Wir wollen nun nur noch mit den Variablen weiterrechnen, die drei oder zwei Gewichtungspunkte erhalten haben, also mit insgesamt 8 Items.

Auf der Basis von Zweier-KFAn über alle $41 + 41 + 41 + 39 + 19 = 181$ Messungen sowohl mit trichotomisierten (i. e. unter Berücksichtigung der Missing Values und der Teilung am Gesamtmedian) wie mit mediandichotomisierten Daten führten wir eine Linkage Analyse durch. Dabei legten wir das Niveau der Signifikanzzahl nach vollständiger Alpha-Adjustierung als Assoziationsmaß zugrunde, bei Gleichheit die Höhe des Prägnanzkoeffizienten.

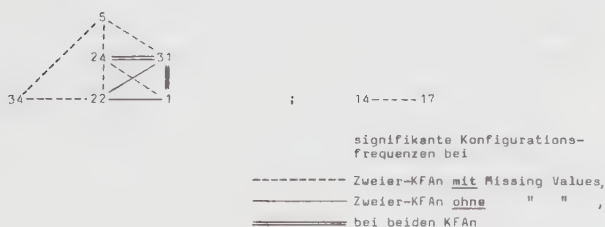


Abbildung 22: Kernstrategiecluster der Linkage Analyse

In Abbildung 22 haben wir die erhaltenen Cluster dargestellt. Danach ist es möglich, die Items 1, 22, 24 und 31 als Kerncluster zu bezeichnen, dem die

Items 5 und 34 angehängt werden können. Die KFAn mit Missing Values werden mit trichotomisierten Daten (über/unter dem Gesamtmedian, Alternativantwort) errechnet; in den KFAn ohne Missing Values wurden die Missing Values aus den Berechnungen ausgeschlossen.

Die Items 14 und 17 stehen für sich. Wir wollen sie von jetzt ab aus arbeitsökonomischen Gründen nicht mehr berücksichtigen.

Für den nächsten Selektionsschritt ist nun eine Überlegung wichtig, die wir des besseren Verständnisses wegen kurz rekapitulieren wollen. Es ist möglich, daß es Prüfungskandidaten mit Strategien gibt und solche ohne! Während wir von ersteren im Idealfall ein stabiles Antwortmuster erhalten, liefern letztere nur Zufallszahlen. Wenn wir die Kandidaten mit Strategien vorher nicht identifizieren, dann werden die Kandidaten ohne Strategien und ihre Zufallszahlen eine Analyse erschweren (die enttäuschenden Ergebnisse der R-Analysen sind möglicherweise eine Folge dieses Zustands).

Aus den verbliebenen 8 Markier-Items müssen wir deshalb diejenigen Items herausfinden, die von einer mehr oder minder großen Gruppe reliabel beantwortet werden; wir untersuchen daher im nächsten Analyseschritt die Retestreliabilität der Markier-Items.

Die Analyse der Differenzwerte über die Datenerhebungszeitpunkte BL, P1, P2, P3 und P4 mit N = 19 und über die Datenerhebungszeitpunkte BL, P1, P2 und P3 mit N = 39 ergab weder für trichotomisierte (kleiner, gleich, größer) noch für dichotomisierte

Daten (kleiner, größer) in den Analysen der einzelnen Items und aller möglichen Zweierkonfigurationen sowohl unter Einbeziehung als auch ohne die Variable Prüfungsart (VD=1/HD=2) einen signifikanten Typ. Ebensovienig fanden wir einen signifikanten Typ bei der Analyse aller Items für je einen der möglichen Differenzwerte. Mit anderen Worten: die Differenzwerte ergeben (anders als z. B. die Differenzen der Emotionstests) unter keinen Umständen ein signifikantes Ergebnis.

Tabelle 51: Ergebnisse der KFAn mit wiederholt gemessenen einzelnen Items

Variablen in der KFA	N	signif. Typ	F	ε	B	s.	Q	s.
Item 1 (BL,P1,P2)	41	222	5	1.13	.520-02	x	.0971	xxx
(BL,P1,P2,P3)	39	2222	3	0.22	.560-02	x	.0717	xxx
(BL,P1,P2,P3,P4)	19	22222	2	0.01	.540-04	xx	.1048	xxx
Item 5 (BL,P1,P2)	41	222	10	3.74	.320-02	x	.1679	xx
(BL,P1,P2,P3)	39	2222	7	1.10	.100-03	xx	.1556	xxx
(BL,P1,P2,P3,P4)	19	22222	5	0.45	.680-04	xx	.2451	xxx
Item22 (BL,P1,P2)	41	-	-	-	-	-	-	-
(BL,P1,P2,P3)	39	2222	6	0.96	.370-03	xx	.1324	xxx
(BL,P1,P2,P3,P4)	19	22222	3	0.09	.110-03	xx	.1536	xxx
Item24 (BL,P1,P2)	41	-	-	-	-	-	-	-
(BL,P1,P2,P3)	39	-	-	-	-	-	-	-
(BL,P1,P2,P3,P4)	19	-	-	-	-	-	-	-
Item31 (BL,P1,P2)	41	111	6	1.03	.530-03	xx	.1243	xxx
(BL,P1,P2,P3)	39	1111	4	0.25	.110-03	xx	.0969	xxx
		2222	20	9.83	.420-03	xx	.3485	xxx
(BL,P1,P2,P3,P4)	19	11111	7	1.75	.100-02	x	.3043	xxx
		22222	3	0.15	.410-03	x	.1513	xxx
Item34 (BL,P1,P2)	41	-	-	-	-	-	-	-
(BL,P1,P2,P3)	39	2222	13	5.35	.150-02	x	.2273	xx
(BL,P1,P2,P3,P4)	19	11111	9	1.74	.180-04	xxx	.4207	xxx
		22222	4	0.14	.120-04	xxx	.2045	xxx
Item 1 (BL-P2)/Prüf.art	41	2222	3	0.33	.450-02	-	.0556	xxx
(BL-P3)/	39	22222	2	0.07	.210-02	-	.0496	xxx
(BL-P4)/	19	222222	2	0.01	.120-04	xxx	.1050	xxx
Item 5 (BL-P2)/	41	-	-	-	-	-	-	-
(BL-P3)/	39	22221	5	0.76	.960-03	x	.1108	xxx
(BL-P4)/	19	-	-	-	-	-	-	-
Item22 (BL-P2)/	41	-	-	-	-	-	-	-
(BL-P3)/	39	22221	5	0.67	.520-03	x	.1130	xxx
(BL-P4)/	19	222221	2	0.05	.110-02	-	.1029	xxx
Item24 (BL-P2)/	41	-	-	-	-	-	-	-
(BL-P3)/	39	-	-	-	-	-	-	-
(BL-P4)/	19	211111	2	0.05	.140-02	-	.1027	xxx
Item31 (BL-P2)/	41	1111	4	0.73	.590-02	-	.0813	xxx
(BL-P3)/	39	11112	2	0.08	.270-02	-	.0494	xxx
(BL-P4)/	19	-	-	-	-	-	-	-
Item34 (BL-P2)/	41	2121	0	6.56	.780-03	x	.1900	x
(BL-P3)/	39	-	-	-	-	-	-	-
(BL-P4)/	19	111112	5	0.82	.110-02	-	.2297	xxx
		222221	3	0.08	.590-04	xx	.1545	xxx

Die Analysen auf der Basis der Rohwerte sind in Tabelle 51 wiedergegeben. Zu beachten ist dabei, daß die jeweils kleineren Versuchspersonenzahlen Teilmengen der jeweils größeren sind.

Wir fanden für die Items 1, 5, 22, 31 und 34 jeweils Kollektive von Prüflingen, die bei einer einmal gewählten Strategie verharren. Lediglich das Item 24 ("Prüfer stellen nur scheinbar leichte Fragen, in Wirklichkeit wollen sie mehr wissen") entzieht sich der formalen Reliabilitätsbestimmung.

Was beim Item 31 und beim Item 34 offensichtlich ist, trifft auch für die anderen Items zu. Während bei ersteren jeweils Typ und Gegentyp signifikant sind, fallen bei letzterem die jeweiligen Gegentypen in das Vertrauensintervall der signifikanten Typen.

Bei Hinzufügung der Variable Prüfungsart fallen - mit einer Ausnahme - die Prägnanzkoeffizienten. Die Ausnahme ist das Item 1, wo wir zwei Hauptdiplomanden finden, die über alle fünf Datenerhebungszeitpunkte das Item "Wenn ich anderer Meinung bin, dann werde ich auch in einer Prüfung einem Professor widersprechen!" konsequent mit "eher falsch" oder "falsch" beantworten.

In Tabelle 52 sind die Analysen von je zwei Items des Strategiekernclusters wiedergegeben. Es zeigt sich, daß die Items 1 und 22 gleichsinnig miteinander kovariieren, beide gegensinnig mit dem Item 31. Wiederum entzieht sich das Item 24 der multiplen Analyse.

Tabelle 52: Ergebnisse der KFA mit je zwei
wiederholt gemessenen SS-Items

Variablen in der KFA	N	signif. Typ	F	E	B	s.	Q	s.
Item 1 (BL,P1,P2)/ Item 22 (BL,P1,P2) <u>41</u>		111 111	9	2.52	.720-03 x		.1683 xxx	
Item 1 (BL,P1,P2,P3,P4)/ Item 22 (BL,P1,P2,P3,P4) <u>19</u>		2222222222	1	0.00005	.540-04 -		.0526 xxx	
Item 1 (BL,P1,P2)/ Item 24 (BL,P1,P2) <u>41</u>		-						
Item 1 (BL,P1,P2,P3,P4)/ Item 24 (BL,P1,P2,P3,P4) <u>19</u>		-						
Item 1 (BL,P1,P2)/ Item 31 (BL,P1,P2) <u>41</u>		222 111	2	0.03	.380-03 x		.0706 xxx	
Item 1 (BL,P1,P2,P3,P4)/ Item 31 (BL,P1,P2,P3,P4) <u>19</u>		2222211111	2	0.001	.460-06 xxx		.1052 xxx	
Item 22 (BL,P1,P2)/ Item 24 (BL,P1,P2) <u>41</u>		-						
Item 22 (BL,P1,P2,P3,P4)/ Item 24 (BL,P1,P2,P3,P4) <u>19</u>		-						
Item 21 (BL,P1,P2)/ Item 31 (BL,P1,P2) <u>41</u>		111 222	9	2.76	.140-02 -		.1211 xxx	
Item 21 (BL,P1,P2,P3,P4)/ Item 31 (BL,P1,P2,P3,P4) <u>19</u>		2222211111	2	0.01	.360-04 x		.1047 xxx	
Item 24 (BL,P1,P2)/ Item 31 (BL,P1,P2) <u>41</u>		222 111	3	0.18	.770-03 x		.0549 xxx	
Item 24 (BL,P1,P2,P3,P4)/ Item 31 (BL,P1,P2,P3,P4) <u>19</u>		-						

Reliable Kernstrategien können wir somit nur von den Items 1, 22 und 31 erwarten. In Tabelle 53 haben wir die Ergebnisse der KFAn dieser drei Items dargestellt. Sie bezieht sich jetzt nur noch auf die Datenerhebungszeitpunkte, bei denen Prüfungen stattfanden; Tabelle 53 wird die Basis für die Agglutination mit den anderen Datensätzen sein.

Tabelle 53: Ergebnisse der KFAn mit den mediandi-
chotomisierten Items der Kernstrategien

variablen in der KFA	Datenerh. zeitpkt.	N	signif. Tyc	F	E	S	s.	C	s.
Item 1, Item 22, Item 31	P1 - P4	140	112 221	42 17	26,89 6,38	.130-02 .220-02	x xx	.140-02 .080-02	xx xxx
	P1	41	-						
	P2	41	-						
	P3	39	-						
	P4	19	-						
Item 1, Item 22, Item 31, Item 32	P1 - P4	140	112 221	42 12	17,84 4,24	.140-02 .140-02	- v	.080-02 .060-02	xx xxx
	P1	41	-						
	P2	41	-						
	P3	39	-						
	P4	19	-						

Wir finden ein starkes Kollektiv, das ggf. einem Prüfer widersprechen will, davon ausgeht, daß ihm eine Antwort auch auf schwere Fragen einfallen wird und es für falsch hält, nicht allzu kritisch aufzutreten. Der weniger prägnante Gegentyp hält es für falsch, einem Professor zu widersprechen, geht nicht davon aus, daß ihm schon etwas einfallen wird und hält es für besser, nicht allzu kritisch aufzutreten. Zum Zwecke der sprachlichen Prägnanz wollen wir künftig erstere als die 'Assertiven', letztere als die 'Unsicheren' bezeichnen.

Die Hinzufügung der Variablen Prüfungsart bringt für beide Typen einen Prägnanzverlust mit sich, den Test gegen die Nullhypothese überleben nur die assertiven und die unsicheren VDn.

Die drei Items 1, 22 und 31 wollen wir im weiteren als Kernstrategien verrechnen.

4.3.3.1.3.3 Selektion von 'Trivialitems' und Analyse der Strukturiertheit

Bei der Strategiensammlung haben wir die Antwortmöglichkeit "noch nie darüber nachgedacht" als Alternative zur 'richtig'-'falsch'-Skala zugelassen. Wir hatten die Summe aller Alternativantworten als Indikator für die Strukturiertheit von Strategien vorgesehen.

Im Anhang haben wir die Mittelwerte, Standardabweichungen und Missing Values für die 34 SS-Items aufgeführt und unter 'MV' die mittlere Häufigkeit, mit der die Versuchspersonen im gesamten Fragebogen die Alternativantwort wählten.

Es zeigt sich, daß im Verlauf der Prüfungsserie die Spalte 'noch nie darüber nachgedacht' weniger und weniger angekreuzt wird. Der Mittelwert des Datenerhebungszeitpunktes P3 fällt dabei aus dem Rahmen. Dieser 'Ausreißerwert' entstand durch die Einbeziehung von zwei Versuchspersonen, die bei P3 ihre letzte Prüfung ablegten und im Interview erklärten, sie hätten über ihre zukünftigen Prüfungen

und alles, was damit zusammenhänge, noch nie nachgedacht. Folglich kreuzten sie bei fast allen Items zum Zeitpunkt P3 die Alternativantwort an. Korrigiert man den Wert für P3, so erhält man für die Häufigkeit der Alternativantworten die Werte $M(BL) = 2.94$, $M(P1) = 0.85$, $M(P2) = 0.55$, $M(P3) = 0.46$ und $M(P4) = 0.35$! Die Alternativantwort wird im Durchschnitt immer seltener gewählt.

Dabei zeigt jedoch ein Blick in die Werte der Items 1 bis 34, daß ein Übergehen von einer unterschiedenen Antwort ('falsch' bis 'richtig') auch während der laufenden Prüfung immer noch möglich ist. Ein pauschaler Strukturiertheitsindex erscheint uns unter diesen Verhältnissen als wenig geeignet, differenziertere Aussagen zu ermöglichen.

Ein anderer Weg, einen Indikator für 'Strukturiertheit' zu gewinnen, geht nicht über die Analyse von Alternativantworten, sondern über die 'Trivialitems'. Als Trivialitem bezeichnen wir diejenigen Items, die von allen Versuchspersonen gleichsinnig beantwortet werden und somit Selbstverständlichkeiten zu erfassen scheinen. Wenn wir davon ausgehen, daß die Selbstverständlichkeiten eine Reflexion des Prüfungssystems darstellen, dann können wir über die Selbstverständlichkeiten diejenigen Kandidaten erfassen, denen wir mangelnde Reflexion unterstellen müssen.

Alle Items, deren Mittelwerte in den Datenerhebungszeitpunkten konstant kleiner/größer als die vorgegebene Skalenmitte von 3.0 zuzüglich oder abzüglich der empirischen Standardabweichungen sind, betrach-

ten wir als von allen Versuchspersonen beantwortet. In diesem Sinne 'selbstverständliche' Items sind (in Gedankenstrichen die 'selbstverständliche' Antwortrichtung):

Item 4: - 'falsch' - ("Ich werde möglichst wenig sagen"),

Item 8: - 'richtig' - ("Am Anfang bin ich sicherlich aufgeregt, aber das wird sich dann legen"),

Item 26: - 'richtig' - ("Es kommt darauf an, Zusammenhänge zu sehen und zu beschreiben"),

Item 28: - 'richtig' - ("Es kommt darauf an, daß ich die Zusammenhänge und Unterschiede anschaulich darstelle") und

Item 18: - 'falsch' - ("Wenn ein Prüfer mir eine Frage stellt, deren Antwort ich überhaupt nicht kenne, dann prüft er mich auf 'sehr gut').

Item 18 nimmt dabei insofern eine Sonderstellung ein, als es in den Datenerhebungszeitpunkten BL und P1 zugleich signifikant selten kalkuliert wird. Vereinfacht gesagt: Die Versuchspersonen haben entweder 'noch nie darüber nachgedacht' oder sie lehnen die Aussage als 'falsch' ab. Möglicherweise kognizieren sie dieses Item als unsinnig.

Wir wollen uns deshalb bei den weiteren Analysen auf die Items 4, 8, 26 und 28 beschränken.

Tabelle 54: Ergebnisse der KFAn mit den median-dichotomisierten SS-Trivialitems

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. Typ	F	E	S	s.	Q	s.
Item 3, Item 26, Item 28	P1 - P4	140	111 212 221 222	8 7 8 69	2,64 15,40 18,08 52,19	.530-o2 .110-o1 .430-o2 .220-o2	x - x x	.0402 xxx .0697 x .0855 x .2007 x	
Item 9, Item 26, Item 28, Prüf.ort	P1 - P4	140	1221 2211 2221	7 4 50	17,45 12,25 35,30	.260-o2 .490-o2 .350-o2	x - -	.0882 x .0665 x .1460 x	

Wir berechneten zwei Typen und zwei Antitypen (i. e. Konfigurationen, die überzufällig selten sind). Der Typ '111' geht aus Messungen hervor, in denen die Items 8 (Am Anfang bin ich sicherlich aufgeregt ...), 26 (... Zusammenhänge sehen und beschreiben) und 28 (... Zusammenhänge und Unterschiede anschaulich darstellen) eindeutig mit "richtig" beantwortet werden; mit ihm werden Merkmalsträger erfaßt, die eindeutig auf Items reagieren, die über die Parameter der Gesamtgruppe als 'trivial' klassifiziert wurden; wir nennen diesen Typ daher die 'Konformen'. Der Typ '222' beruht auf Messungen, in denen diese drei Trivialitems nicht eindeutig ("eher richtig" oder "starkes Problem für mich, weiß ich aber nicht" oder "eher falsch" oder "falsch") skaliert werden; mit ihm werden Merkmalsträger erfaßt, die nicht eindeutig antworten, obwohl diese Items über die Parameter der Gesamtgruppe als trivial klassifiziert wurden; zum Zwecke semantischer Prägnanz nennen wir diesen Typ die 'Differenzierten'.

Die beiden Antitypen zeigen, daß in unseren Messungen eine eindeutige Antwort auf das Item 26 oder das Item 28 in Kombination mit nichteindeutigen Antworten auf die beiden anderen Items überzufällig selten auftritt.

In der unteren Hälfte der Tabelle sind die Ergebnisse eines Wiederholungslaufes unter Hinzufügung der Variablen Prüfungsart dargestellt. Es zeigt sich, daß die 'Konformen' nunmehr die Signifikanzgrenze verfehlen und daß die 'Differenzierten' unter Prägnanzverlusten der Klasse der VDN zugeordnet werden.

Laufwiederholungen - dieses Mal für jeden Datenerhebungszeitpunkt getrennt - ergaben, daß in keinem der Läufe eine Konfigurationsfrequenz das alpha-adjustierte Signifikanzniveau von 5 % erreichte.

Ferner überprüften wir die formale Reliabilität der Rohwerte. Hier erhielten wir einen signifikanten Typ mit der Merkmalsausprägung "2" (i. e. nicht-eindeutige Antworten) in allen Analysen. Unter leichten Prägnanzverlusten wurden die Typen dem Kollektiv der VDN zugeordnet. Wir haben dieses Ergebnis nicht tabellarisch dargestellt, weil es aus Tabelle 54 unter dem Gesichtspunkt hoher Wahrscheinlichkeit hervorgeht.

4.3.3.2 Zusammenfassung und Überblick

Bis hierhin haben wir jeden Datensatz zur Messung der Strategien einzeln für sich überprüft.

Zur Messung der Strategien verwendeten wir vier Datensätze - TVD(W), TVD(H), Kernstrategieitems aus der SS und Trivialitems aus der SS - sowie die Taktiken vom Beisitzerblatt.

Bei der Einzelanalyse der Kernstrategien hatten wir die Gegentypen der 'Unsicheren' und der 'Assertiven' gefunden; die Hinzufügung der Variablen 'Prüfungsart' bringt für beide Typen Prägnanzverluste mit sich, beide Typen werden der Klasse der VDN zugeordnet.

Bei der Analyse der Trivialitems fanden wir die Typen der 'Konformen' und der 'Differenzierten'; bei Hinzufügen der Variablen 'Prüfungsart' werden die 'Konformen' aufgespalten und verfehlen nunmehr die Signifikanzgrenze, werden die 'Differenzierten' unter Prägnanzverlusten der Klasse der VDN zugeordnet, sind nunmehr lediglich auf dem 10-%-Niveau signifikant. Bei den Einzelanalysen der TVD-Daten errechnen wir bei den prospektiv skalierten Daten des TVD(W) zwei Gegentypen: einen Typ, der in Zukunft mäßig-angespannt, variabel und diversifizierend vorgehen will, und einen Gegentyp der in Zukunft angespannt, stetig und unifizierend vorgehen möchte; bei Hinzufügen der Variablen Prüfungsart wird letztere aufgespalten und verfehlt nunmehr die Signifikanzgrenze, ersterer wird der Klasse der VDN zugeordnet.

Bei der Einzelanalyse der TVD(H)-Daten errechnen wir erst nach Hinzufügen der Variablen Prüfungsart einen Hinweis auf ein Kollektiv von HDn, das angespannt, stetig und unifizierend vorgegangen ist. Bei der gemeinsamen Analyse von TVD(W)- und TVD(H)-Daten zeigte sich, daß intendierte und realisierte Strategien die gleiche Struktur haben; bei Hinzufügen der Variablen Prüfungsart wird der Typ derjenigen, die mäßig-angespannt, variabel und diversifizierend vorgegangen ist und weiterhin vorgehen will, der Klasse der VDn zugeordnet; der Typ derjenigen, die angespannt, stetig und unifizierend vorgegangen ist und vorgehen möchte, wird beiden Anthac-Klassen zugeordnet.

Agglutinations-KFAn aller zur Messung der Strategien eingesetzten Datensätze brachten überwiegend Ergebnisse, in denen der Prägnanzkoeffizient um mehr als 50 % sank oder gar kein signifikanter Typ mehr berechnet werden konnte. Wir wollen diese Ergebnisse daher nicht en detail aufführen, sondern lediglich die Ausnahmen in verbaler Formulierung:

- Beide von uns klassifizierten Kollektive mit reliablen Kernstrategien - die 'Assertiven' und die 'Unsicheren' - lassen die Prüfer überzufällig häufig ausreden, i. e. beide Kollektive werden überzufällig selten unterbrochen.
- Für die Datensätze TVD(W), TVD(H) und Kernstrategien können wir ein in allen Analysen wiederkehrendes Ergebnis verallgemeinern: Das Kollektiv der 'Assertiven' intendiert und realisiert ein angespanntes, stetiges und unifizierendes Vorgehen,

der Gegentyp der Unsicheren intendiert und realisiert ein mäßig angespanntes, variables und diversifizierendes Vorgehen. In den Wiederholungsläufen unter Hinzufügung der Variablen Prüfungsart wird der ersterwähnte Typ der Anthac-Klasse der HDn, der zweiterwähnte der Anthac-Klasse der VDn zugeordnet.

Hypothese 9 lautete: "Es gibt mindestens zwei multivariate Profile, wovon das eine einer Kontrollstrategie, das andere der Strategie der Gelernten Hilflosigkeit entspricht und wovon das eine die Umkehrung des anderen ist." In allen verwendeten Datensätzen fanden wir (mit Ausnahme der Taktiken) zwei gegenläufige multivariate Profile: den 'mäßig angespannt-variabel-diversifizierend' Vorgehenden stehen die 'angespannt-stetig-unifizierend' Vorgehenden gegenüber, die 'Assertiven' bilden den Gegentyp zu den 'Unsicheren', zudem kovariieren die jeweils erstgenannten Typen einerseits, die jeweils letztgenannten Typen andererseits; bei den Trivialitems fanden wir die 'Konformen' und die 'Differenzierten'. Der formale Teil der Hypothese - die Existenz zweier gegenläufiger, multivariater Profile - kann daher für alle Datensätze angenommen werden.

Inhaltlich können wir bei den 'Assertiven' (die zudem angespannt, stetig und unifizierend vorgegangen sind und vorgehen wollen) ein aktiv eingreifendes und steuerndes Verhalten als Strategie annehmen, bei den 'Unsicheren' (die zudem mäßig angespannt, variabel und diversifizierend sich verhalten wollen und sich verhalten haben) ein passiv,

treiben-lassendes Verhalten. Unterschiede in den Kontrollaspekten, d. h. in der Bereitschaft, ggf. aktiv in das Prüfungsgeschehen einzugreifen und von sich aus den Prüfungsverlauf zu beeinflussen, sind deutlich zu erkennen.

Die Trivialitems messen einen anderen Aspekt als die anderen Datensätze. Inhaltlich liegt ihnen mehr eine Kalkulation von diskriminativen Reizen zugrunde, wobei die 'Konformen' diese Diskrimination geleistet haben, die 'Differenzierten' dabei unsicher sind. Auch diesen Aspekt können wir theoretisch auf die Gelernte Hilflosigkeit beziehen (vgl. S. 96 ff). Der inhaltliche Aspekt (der Gegensatz zwischen Kontrollstrategie und Gelernter Hilflosigkeit stellt sich hier als unterschiedliche Diskriminationsleistung dar.

In den Hypothesen 3 und 5 hatten wir postuliert, daß sich VDN und HDN in ihren Strategien und Verhaltensweisen unterscheiden, und zwar daß VDN die Strategie der Gelernten Hilflosigkeit eher realisieren als Kandidaten in Abschlußprüfungen. Formal entspricht dies einer Prognose der Assoziation zwischen der Merkmalsausprägung 'VD' der Variablen Prüfungsart und dem Profil, das wir als 'Gelernte Hilflosigkeit' bezeichnet haben, sowie der Prognose einer Assoziation zwischen der Merkmalsausprägung 'HD' und dem 'Kontrollstrategie'-Profil.

Mit Ausnahme der Ergebnisse der Kernstrategieitems ergaben sich in den einzelnen Datensätzen nach Hinzufügen der Variablen Prüfungsart strukturell gleiche

Ergebnisse (wie schon bei den Analysen der Belastungsmessungen): die Kollektive der 'Differenzierten' und der 'mäßig angespannt-variabel-diversifizierend Vorgehenden' werden der Merkmalsausprägung VD zugeordnet, während die jeweiligen Gegentypen aufgespalten werden und nunmehr z. T. die Signifikanzgrenze verfehlen. Nach dem strengen Algorithmus annehmen können wir die Hypothese lediglich für den aus TVD(W) und TVD(H) zusammengesetzten Datensatz. Aufgrund der Ergebnisse der Kernstrategieitem müssen wir die in den Hypothesen enthaltene Kontrastbehauptung explizit ablehnen.

Im folgenden wollen wir die Ergebnisse darstellen, die sich auf die Hypothesen 4 und 7 beziehen, also auf die Zusammenhänge zwischen Strategien und Leistungsdaten einerseits sowie zwischen Strategien und Prüfungsbelastungen andererseits.

4.3.4 Strategien und Belastungen (Hypothese 7, Hypothesentriade zu Vorerfahrungen, Prüfungsstrategien und Prüfungsbelastungen)

Hypothese 7 lautete: "Kandidaten, die die Strategie der Gelernten Hilflosigkeit realisieren, sind stärkeren Prüfungsbelastungen unterworfen als Kandidaten, die die Kontrollstrategie einsetzen." In der Hypothesentriade haben wir unter Anwendung des Immunisierungsparadigmas postuliert, daß Vorerfahrungen mit Prüfungen zu geringeren Prüfungsbelastungen führen und zum Einsatz von Kontrollstrategien führen, bzw. umgekehrt: daß mangelnde Vorerfahrungen mit überdurchschnittlichen Belastungen und Gelernter Hilflosigkeit einhergehen.

Im folgenden stellen wir die Ergebnisse der Agglutinations-KFAn vor, in denen die Strategie-Datensätze (Kernstrategien, Trivialitems, Test-verhaltensdifferentiale) und die Belastungs-messungen verrechnet wurden. Die jeweils untere Hälfte der Tabelle gibt die Ergebnisse von Wiederholungsläufen unter Hinzufügung der Variablen 'Prüfungsart' wieder.

In Tabelle 55 sind die Ergebnisse der gemeinsamen Analyse von Kernstrategieitems und Emotionstests dargestellt. Wir berechneten drei signifikante Typen:

Tabelle 55: Ergebnisse der KFAn mit den Kernstrategien und den Emotionstests

variablen in c r f a	datennr. stafordt.	N	signif. typ	r	s	r	s.	s	s.
Item 1, Item 20, Item 31, His, STAI, BDI	1 - 14	140	112111 112122 221111 -	1 13 9	1.43 1.43 0.43	1.43=01 xxx 1.43=04 xx 1.43=03 x	1.43 xxx 1.43 xxx 1.43 xxx	1.43 xxx 1.43 xxx 1.43 xxx	1.43 xxx 1.43 xxx 1.43 xxx
	P1	41	-	-	-	-	-	-	-
	P2	41	112111	5	0.46	1.43=04 xx	1.43 xxx	1.43 xxx	1.43 xxx
	P3	39	-	-	-	-	-	-	-
	P4	19	221211	3	0.15	1.43=03 x	1.43 xxx	1.43 xxx	1.43 xxx
K.D., Prof.ert	1 - 14	140	112111 112222	11 11	1.43 1.43	1.43=04 xx 1.43=04 xx	1.43 xxx 1.43 xxx	1.43 xxx 1.43 xxx	1.43 xxx 1.43 xxx
	P1	51	-	-	-	-	-	-	-
	P2	41	-	-	-	-	-	-	-
	P3	39	-	-	-	-	-	-	-
	P4	19	-	-	-	-	-	-	-

Wir finden einen Typ mit 'assertiven' Kernstrategien und niedrigen emotionalen Reaktionen, einen Typ mit 'assertiven' Kernstrategien und hohen emotionalen Reaktionen und einen - schwach prägnanten - Typ mit 'unsicheren' Kernstrategien und unterdurchschnittlichen Scores im STAI und überdurchschnittlichen in His und BDI.

Die Hinzufügung der Variablen Prüfungsart ergibt, daß VDN mit 'assertiven' Kernstrategien entweder unterdurchschnittliche Scores in den Emotionstests oder überdurchschnittliche Scores skalieren.

Ein Vergleich mit den Einzelanalysen zeigt, daß wir lediglich beim Kollektiv der Assertiv-Niedrig-emotionalisierten VDN keine drastischen Pränanzverluste registrieren müssen. Eine Assoziation zwischen Kernstrategien und den Emotionstests kann auf der Basis der bisherigen Ergebnisse nur für dieses Kollektiv gelten.

Auffallend und interessant sind hier mehr andere Ergebnisse: Das aus der Einzelanalyse der Emotions-tests bekannte Kollektiv der Hoch-Emotionalisierten ist besonders drastischen Pränanzverlusten unterworfen, und die Klasse der HDn kann sich in keiner der signifikanten Konfigurationen durchsetzen.

Die gemeinsame Analyse der Kernstrategien und der Analogskalen 'Erschöpfung' und 'Depression' ergab lediglich für den Lauf über alle vier Datenerhebungszeitpunkte signifikante Konfigurationen. Sie sind in Tabelle 56 dargestellt.

Tabelle 56: Ergebnisse der KFA mit den Kernstrategien und den Analogskalen des Fragebogens AN

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. Typ	F	E	3	s.	1	s.
Item 1, Item 22, Item 31, Erschöpfung, Konzentration	P1 - F4	133	11222 22122	14 6	5.75 1.30	.1910-n3	x	.1701 xxx	
						.1110-c2	-	.13700 xxx	
s.o., Prüf.art	P1 - F4	133	-						

Gegenüber beiden Einzelanalysen ergibt sich ein drastischer Prägnanzverlust, der im Falle der Assertiven-Erschöpften allerdings weniger als 50 % beträgt.

Die gefundenen Typen deuten auf eine gewisse Assoziation zwischen niedriger Konzentrationsfähigkeit und starker Erschöpfung mit intendiertem 'assertivem' oder 'unsicherem' Vorgehen hin, ohne daß ein Bezug zur Variable Prüfungsart festzustellen wäre.

Keine der gemeinsamen Analysen der Kernstrategien und der Faktorenwerte der Beschwerdeliste - sowohl über alle Daten wie in Einzelanalysen über die einzelnen Datenerhebungszeitpunkte - brachte eine signifikante Konfigurationsfrequenz, auch nicht nach Hinzufügen der Variable Prüfungsart.

Insgesamt gesehen ergibt sich somit folgendes Ergebnis: Prospektiv skalierte 'assertive' Kernstrategien gehen mit niedrigen oder starken Emotionen während der Prüfung einher; über 'unsichere' Kernstrategien kann keine Aussage gemacht werden. Sowohl 'assertive' als auch 'unsichere' Kernstrategien führen zu überdurchschnittlicher 'Erschöpfung' und 'Konzentrationsunfähigkeit' nach der Prüfung. Nun wissen wir aber, daß im Falle der Kernstrategien wir es mit Minikollektiven mit stabilen Strategien zu tun haben, denen logischerweise eine Gruppe mit instabilen Strategien gegenüberstehen muß. Auf der Basis der bislang dargestellten Ergebnisse können wir lediglich vermuten, daß

instabile Strategien etwas mit dem 'Verschwinden' des hochprgnanten Kollektivs derjenigen zu tun hat, die nach der Prfung niedrige Erschpfung und hohe Konzentrationsfhigkeit skalieren.

In Tabelle 57 sind die Ergebnisse der gemeinsamen Analyse der Strategietrivialitems und der Emotionstests ber alle Datenerhebungszeitpunkte dargestellt sowie des Wiederholungslaufes unter Hinzufgung der Variablen Prfungsart (VD=1)

Tabelle 57: Ergebnisse der KFAn mit den Strategietrivialitems und den Emotionstests

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. Typ	F	E	B	s.	Q	s.
Item 8,Item 26,Item 28, HiS,STAI,BOI	P1 - P4	140	222222	27	6.64	.450-o9	xxx	.1575	xxx
Item 8,Item 26,Item 28, HiS,STAI,BOI,Prf.art	P1 - P4	140	.1121112 2222221	4 19	o.3o 4.48	.270-o3 .130-o6	x xxx	.o272 .11o4	xxx xxx

Wir finden ein einziges hochsignifikantes prgnantes Kollektiv, in dem eine nicht eindeutige Beantwortung der Trivialitems mit berdurchschnittlichen Emotionen einhergeht. Gegenber beiden Einzelanalysen sind die Prgnanzverluste relativ gering.

Bei Hinzufgung der Variablen Prfungsart (VD=1) zeigt sich, da 70 % des o. e. Kollektivs sich aus der Gruppe der VDN rekrutiert, wobei ein weiterer Prgnanzverlust hingenommen werden mu. Gleichzeitig finden wir einen neuen signifikanten, schwach prgnanten Typ, der ein Kollektiv von HDN darstellt, die mit unterdurchschnittlichen Emotionen Item 8 und Item 26 eindeutig beantworten, bei Item 28 nicht eindeutige Reaktionen zeigen.

In Tabelle 58 sind die Ergebnisse der gemeinsamen Analyse über alle Datenerhebungszeitpunkte für die Strategietrivialitems und die Faktorscores der Beschwerdeliste sowie eines Wiederholungslaufes unter Hinzufügung der Variablen Prüfungsart dargestellt.

Tabelle 58: Ergebnisse der KFA mit den Strategietrivialitems und Faktorscores der Beschwerdeliste

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. Typ	F	E	S	s.	Q	s.
Item 9,Item 26,Item 28, BL1,BL2	P1 - P4	140	-						
Item 9,Item 26,Item 28, BL1,BL2,Prüf.art	P1 - P4	140	222221	18	6.98	.220-03	x	.0854	xxx

Bei der gemeinsamen Analyse der beiden Datensätze ergibt sich kein signifikanter Typ. Erst nach Hinzufügung der Variablen Prüfungsart finden wir ein signifikantes, mäßig prägnantes Kollektiv von VDn, das die Trivialitems nicht eindeutig beantwortet und in der Beschwerdeliste überdurchschnittliche Werte erreicht.

In Tabelle 59 sind die Ergebnisse der gemeinsamen Analyse über die Datenerhebungszeitpunkte für die Strategietrivialitems und die beiden verbliebenen Analogscores sowie des Wiederholungslaufes unter Hinzufügung der Variablen Prüfungsart dargestellt.

Tabelle 59: Ergebnisse der KFAn mit den Strategie-trivialitems und den Analogskalen des Fragebogens AN

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. Typ	F	E	B	s.	2	s.
Item 8,Item 26,Item 28, Erschöpfung,Konzentration	P1 - P4	131	11111 22222	5 22	0.70 10.16	.760-03 .470-03	x x	.0340 .1013	xxx xxx
Item 8,Item 26,Item 28, Erschöpfung,Konzentration, Prüf.art	P1 - P4	131	111111 222221	5 21	0.51 7.36	.180-03 .140-04	x xxx	.0355 .1140	xxx xxx

Wir finden zwei gegenläufige Typen. Beim einen Kollektiv sind eindeutige Antworten auf die Trivialitems mit unterdurchschnittlicher Erschöpfung und überdurchschnittlicher Konzentrationsfähigkeit assoziiert, beim anderen Kollektiv sind nicht eindeutige Antworten mit überdurchschnittlicher Erschöpfung und unterdurchschnittlicher Konzentrationsfähigkeit assoziiert.

Die Hinzufügung der Variablen Prüfungsart (VD=1) zeigt, daß es sich bei den o. a. Kollektiven um Kollektive von VDn handelt, was sich in gleichbleibender oder wachsender Prägnanz niederschlägt. Die gefundene Assoziation zwischen den Antworten auf die Trivialitems und den Analogskalen gilt demnach nur für das Subkollektiv der VDn.

In allen drei Analysen zeigt sich ein gemeinsames Ergebnis: Die nicht-eindeutigen Antworten auf die drei Trivialitems gehen mit überdurchschnittlichen Emotionen einher. Am deutlichsten zeigt sich dies bei der gemeinsamen Analyse mit den Analogscores,

wo auch ein - wenngleich schwach prägnanter - signifikanter Gegentyp gefunden werden konnte. Dieser Gegentyp, der eine Assoziation zwischen eindeutigen Reaktionen und unterdurchschnittlichen Emotionen markiert, konnte in den beiden anderen gemeinsamen Analysen nicht gesichert werden.

Weiter zeigt sich, daß die Assoziation der nicht eindeutigen Reaktionen und der überdurchschnittlichen Emotionen auf das Subkollektiv der VDN beschränkt ist. Eindeutig ist dies bei der gemeinsamen Analyse mit den Analogskalen. Bei den Emotionstests und der Beschwerdeliste ist dies weniger deutlich.

Zusammenfassend läßt sich sagen, daß innerhalb der Gruppe der VDN differenziertes Antworten auf trieviale Items mit überdurchschnittlichen Belastungen einhergeht.

In Tabelle 60 sind die vier TVD-Subsätze und die drei LIKERT-skalierten Tests HiS, STAI und BDI miteinander verrechnet worden und die Ergebnisse dargestellt. Nur eine einzige der signifikanten Konfigurationsfrequenzen der Tabelle erreicht einen Wert, der größer als $Q=.1000$ ist. Bei der Analyse der LIKERT-Skalen wurde für keine der signifikanten Konfigurationsfrequenzen ein Wert berechnet, der kleiner ist als $Q=.1000$; dies gilt auch für die TVD-Analysen innerhalb der einzelnen Datenerhebungszeitpunkte. Im allgemeinen brachte die Agglutination der Variablensätze also keinen Prägnanzvorteil.

Tabelle 60: Ergebnisse der KFAn mit den Skalen
des TVD und den Emotionstests

Variablen in der KFAn	Intererh. zeitpunkt	n	signif. Typ	F	E	B	S.	...	...
A(J=H), U(H), I(H), HIS, STAI, BDI	P1 - P4 P1 P2 P3 P4	140 41 41 39 19	121111 - - - -	10 4	2,43 0,79	,180-03 ,560-03	x x	,0550 xxx ,1111 xxx	
A(H), U(H), I(H), HIS, STAI, BDI	P1 - P4 P1 P2 P3 P4	140 41 41 39 19	112111 121222 212111 221222 112111 - - -	10 9 9 9 4	2,67 2,18 2,12 1,76 0,36	,320-03 ,380-03 ,110-03 ,000-04 ,470-03	x x x xx x	,0530 xxx ,0430 xxx ,0409 xxx ,0567 xxx ,0366 xxx	
A(J=H), U(H=H), I(H=H), HIS, STAI, BDI	P1 - P4 P1 P2 P3 P4	140 41 41 39 19	122111 - - - -	5	0,43	,780-04	x	,1321 xxx	
A(J), I(J), U(J), A(H), U(H), I(H), HIS, STAI, BDI	P1 - P4 P1 P2 P3 P4	140 41 41 39 19	121121222 21212111 21221211 22212111 121121222 -	6 6 4 4 3 3	0,32 0,22 0,22 0,22 0,10 0,11	,110-05 ,120-01 ,900-04 ,900-04 ,130-03 ,170-03	xxx xxx x x - -	,0407 xxx ,0413 xxx ,0270 xxx ,0270 xxx ,0710 xxx ,1700 xxx	
A(J), I(J), U(J), A(H), U(H), I(H), HIS, STAI, BDI	P1 - P4 P1 P2 P3 P4	140 41 41 39 19	121111 - - - -	5	0,73	,780-03	-	,1111 xxx	
A(J), I(J), U(J), A(H), U(H), I(H), HIS, STAI, BDI	P1 - P4 P1 P2 P3 P4	140 41 41 39 19	112111 - - - -	3	0,11	,170-03	x	,0701 xxx	
A(J), I(J), U(J), A(H), U(H), I(H), HIS, STAI, BDI	P1 - P4 P1 P2 P3 P4	140 41 41 39 19	121111 - - - -	2	0,02	,210-03	-	,1111 xxx	
A(J), I(J), U(J), A(H), U(H), I(H), HIS, STAI, BDI	P1 - P4 P1 P2 P3 P4	140 41 41 39 19	12112221 - - - -	3	0,07	,510-04	-	,1111 xxx	

Bei der näheren Analyse zeigt sich jedoch, daß die
TVD(Werde)-Daten, die bei der Einzelanalyse (a)
überhaupt Typen lieferte und (b) sich als reliabel

erwiesen, unter dem Pränanzverlust besonders leiden. Die TVD(Habe)-Daten liefern plötzlich Typen, bei der Analyse über alle Datenerhebungszeitpunkte sogar vier Stück. Durch Agglutination dieser Typen gelangt man zur Aussage, daß realisiertes stetiges und global-synthetisierendes Vorgehen mit niedrigen Emotionen einhergeht, während variables und diversifizierend-analytisches Vorgehen mit hohen Emotionen einhergeht. Von den betroffenen Versuchspersonen wollen dabei zwei Drittel auch künftig so vorgehen. Weiterhin zeigt sich, daß von den insgesamt 14 von 140 Merkmalsträgern, die angespannt, stetig und global vorgegangen sind und dieses weiter tun wollen, 71 % niedrige Ausprägungen in den Emotionstests HiS und BDI erreichen.

In Tabelle 61 sind die Ergebnisse der KFA für die TVD-Daten und die Analogskalen Erschöpfung (2 = völlig) und Konzentration (2 = niedrig) dargestellt. Innerhalb der einzelnen Datenerhebungszeitpunkte ergab sich keinerlei signifikante Konfigurationsfrequenz. Die in Tabelle 61 angegebenen Pränanzkoeffizienten liegen alle deutlich unter den aus den Einzelanalysen bekannten. Im allgemeinen ist also ein deutlicher Pränanzverlust eingetreten.

Tabelle 61: Ergebnisse der KFA mit den Skalen des TVD und den Analogskalen des Fragebogens AN

Variablen in der KFA	signif. zeltakt.	N	signif. /N						
A(W), V(W), U(W), Erschöpfung, Konzentration	P1 - P4	131	-						
A(H), V(H), U(H), Erschöpfung, Konzentration	P1 - P4	131	-						
A(W), V(W), U(W), A(H), U(H), U(H), Erschöpfung, Konzentration	P1 - P4	131	21221222 12112121	6 7	0,73 0,6	.110-05 xxx .402-05 xx		.0674 xxx .0447 xxx	
A(W), V(W), U(W), Erschöpfung, Konzentration, Prüf. art	P1 - P4	131	212221 121221	9 10	1,99 1,11	.190-03 x .120-02 -		.0567 xxx .0537 xxx	
A(H), V(H), U(H), Erschöpfung, Konzentration, Prüf. art	P1 - P4	131	-						
A(W), V(W), U(W), A(H), U(H), U(H), Erschöpfung, Konzentration, Prüf. art	P1 - P4	131	212212221 121121211	5 5	0,24 0,46	.470-05 xx .110-03 -		.0554 xxx .0740 xxx	

Lediglich bei zwei Typen ist der Pränanzverlust geringer als 50 %. Es handelt sich zum einen um das Kollektiv derjenigen, die zukünftig angespannt, stetig und unfizierend vorgehen wollen und überdurchschnittliche Werte in den Analogskalen skalieren; verglichen mit der Einzelanalyse der TVD(W)-Daten ist der Pränanzverlust weniger drastisch; der Wiederholungslauf unter Hinzufügung der Variablen Prüfungsart ergibt, daß sich 82 % dieses Kollektivs aus der Klasse der VDN rekrutieren. Zum anderen handelt es sich um das Kollektiv der VDN, deren realisierte und intendierte Strategie als angespannt, stetig und unfizierend beschrieben

werden muß und die überdurchschnittliche Werte in den Analogskalen skalieren.

Bei keiner der gemeinsamen Analysen der Daten des TVD und der Faktorscores der Beschwerdeliste fand sich eine signifikante Konfigurationsfrequenz.

Zusammenfassend finden wir in den gemeinsamen Analysen von TVD-Datensätzen und Belastungsmessungen die folgenden Ergebnisse: Im allgemeinen verzeichnen wir deutliche bis drastische Prägnanzverluste. Die retrospektiv skalierten TVD(Habe)-Daten ergeben, daß realisiertes stetiges und unifizierendes Vorgehen mit niedrigen Emotionen und variables, diversifizierendes Vorgehen mit starken Emotionen während der Prüfung einhergeht. Die prospektiv skalierten TVD(Werde)-Daten ergeben, daß intendiertes, angespanntes, stetiges und unifizierendes Vorgehen bei den VDn mit überdurchschnittlicher Erschöpfung und Konzentrationsunfähigkeit nach der Prüfung einhergeht. In beiden Fällen handelt es sich um Kollektive, bei denen das Profil der intendierten und realisierten Strategien übereinstimmt. In keinem Fall kann ein Zusammenhang von den Ausprägungen der Emotionsvariablen her zu den TVD-Strategien hin hergestellt werden.

Als Hypothese 7 hatten wir formuliert: "Kandidaten, die die Strategie der Gelernten Hilflosigkeit realisieren, sind stärkeren Prüfungsbelastungen unterworfen als Kandidaten, die Kontrollstrategien einsetzen!" Weiterhin haben wir die unterdurchschnittlichen Ausprägungen der Belastungsmessungen und die 'Kontrollstrategie' in der Hypothesentriade zur Klasse der HDn, überdurchschnittliche Belastungen und 'Gelernte Hilflosigkeit' zur Klasse der VDn in Beziehung gesetzt.

Wenn wir die Kernstrategieitems verrechnen, müssen wir Hypothese und Hypothesentriade ablehnen.

Verrechnen wir hingegen die Trivialitems, so finden wir nicht nur die vorhergesagten Assoziationen, sondern auch die implizite Kontrastaussage der Hypothese 7 in der Agglutinations-KFA mit den Analogskalen zur Erfassung der Postprüfungsbelastung voll bestätigt: 'Konforme' sind unterdurchschnittlich belastet, 'Differenzierte' skalieren überdurchschnittliche Merkmalsausprägungen. In den Agglutinations-KFAn mit den Trivialitems einerseits und den beiden anderen Belastungs-Datensätzen zeigen sich jeweils Assoziationen zwischen 'differenzierten' Strategien und überdurchschnittlichen Belastungen; das Kollektiv der 'Konformen' wird aufgespalten und verfehlt jeweils die Signifikanzgrenze (ein Ergebnis, dessen Analogie wir bereits kennengelernt haben).

Wiederholungsläufe für alle drei Agglutinations-KFAn unter Hinzufügung der Variablen Prüfungsart ergeben, daß in allen drei KFAn sich die Assoziationen 'differenzierter' Strategien mit überdurch-

schnittlichen Belastungen und dem Merkmal VD als signifikant erweist. Zumindest den Teil der Hypothesentriade, der sich auf die VDn bezieht können wir annehmen, die implizite Kontrastaussage bezüglich der HDn verfehlt jedoch die Signifikanzgrenze.

Inkonsistent sind die Ergebnisse bei den TVD-Datensätzen: teils berechtigen uns die Ergebnisse zur Annahme der Hypothese 7, teils müssen wir sie ablehnen.

4.3.5 Strategien und Leistungsdaten (Hypothese 4, Hypothesentriade zu Vorerfahrungen, Prüfungsstrategien und Leistungsdaten)

Hypothese 4 lautete: "Erfolgreiche Kandidaten realisieren aktive und kontrollierende Strategien, während nicht-erfolgreiche passiv und hilflos agieren." In der Hypothesentriade haben wir unter Anwendung des Immunisierungsparadigmas postuliert, daß Vorerfahrungen mit Prüfungen mit Kontrollstrategien und überdurchschnittlichen Leistungsdaten einhergehen, während umgekehrt mangelnde Vorerfahrungen zu Gelernter Hilflosigkeit und unterdurchschnittlichen Leistungsdaten führen.

In Tabelle 62 sind die Ergebnisse der gemeinsamen Analyse der Kernstrategien und der Leistungsdaten ('erwartete' und 'erhaltene' Zensur und der skalierte 'Prozentsatz des Stoffes', der im Studium bekannt wurde) aufgeführt.

Tabelle 62: Ergebnisse der KFA mit den Kernstrategien und den Leistungsdaten

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. Typ	F	E	S	+	1	s.
Item 1, Item 22, Item 31, erw.Zensur, Prozentsatz, ern.Zensur	P1 - P4	133	112121 112212	3 15	1.52 5.45	.160-03 .370-03	x x	.051- .0500	xxx xxx
Item 1, Item 22, Item 31, erw.Zensur, Prozentsatz, erh.Zensur, Prüf.art	P1 - P4	133	1212121	11	3.17	.360-03	x	.0559	xxx

Bei den Analysen für die einzelnen Datenerhebungszeitpunkte fand sich keine signifikante Konfigurationsfrequenz.

Wir fanden in der Gesamtanalyse über vier Datenerhebungszeitpunkte zwei Kollektive: 'Assertive' mit überdurchschnittlichen Leistungsdaten und 'Assertive' mit unterdurchschnittlichen Leistungsdaten. Gegenüber beiden Einzelanalysen fällt die Prägnanz drastisch ab.

Bei Hinzufügung der Variablen Prüfungsart finden wir ein mäßig prägnantes Kollektiv von VDn, das dem Prüfer widersprechen will, aber nicht davon ausgeht, daß ihnen ggf. etwas einfallen wird, und auch nicht allzu kritisch auftreten will, dabei unterdurchschnittliche Leistungswerte erreicht.

In Tabelle 63 sind die Ergebnisse der gemeinsamen Analyse über alle Datenerhebungszeitpunkte für die Strategietrivialitems und die Leistungsdaten sowie eines Wiederholungslaufes unter Hinzufügung der Variablen Prüfungsart dargestellt.

Tabelle 63: Ergebnisse der KFA mit den Strategietrivialitems und den Leistungsdaten

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. Typ	F	E	B	s.	Q	s.
Item 9, Item 26, Item 28, erw. Zensur, Prozentsatz, erh. Zensur	P1 - P4	131	111121 121121 222212	3 6 30	0,14 0,49 11,12	.390-03 .110-04 .430-06	x xxx xxx	.0226 .0436 .1629	xxx xxx xxx
Item 9, Item 26, Item 28, erw. Zensur, Prozentsatz, erh. Zensur, Prüf. art	P1 - P4	131	1211211 2222121	4 25	0,35 8,06	.480-03 .390-06	- xxx	.0289 .1424	xxx xxx

Gegenüber beiden Einzelanalysen sind leichte Prägnanzverluste zu beobachten. Wir finden zwei Typen, die sich aus den Subkollektiven der Einzelanalysen zusammensetzen: ein Kollektiv mit eindeutigen Reaktionen auf die Trivialitems und überdurchschnittlichen Leistungsmerkmalen und ein zweites Kollektiv mit nicht-eindeutigen Reaktionen und unterdurchschnittlichen Leistungsdaten. Ferner finden wir ein drittes Kollektiv mit überdurchschnittlichen Leistungsdaten, das auf die Items 8 und 28 eindeutig, nicht eindeutig auf das Item 26 reagiert.

Als Ergebnis eines Wiederholungslaufes unter Hinzufügung der Variablen Prüfungsart erhalten wir nur noch ein hochsignifikantes, zudem prägnantes Kollektiv: 'differenzierte' VDN mit unterdurchschnittlichen Leistungsdaten.

Die Ergebnisse der gemeinsamen Analyse der Trivialitems und der Leistungsdaten lauten verbal formuliert: 'konforme' Strategien und überdurchschnittliche Leistungsdaten treten überzufällig häufig bei

denselben Merkmalsträgern auf, 'differenzierte' Strategien und unterdurchschnittliche Leistungsdaten treten überzufällig häufig bei denselben Merkmals-trägern auf, die zudem überzufällig häufig aus der Klasse der VDN stammen.

In Tabelle 64 finden sich die Ergebnisse der gemeinsamen Analyse der Leistungsdaten und des TVD in beiden Versionen dargestellt.

Tabelle 64: Ergebnisse der KFAn mit den TVD-Scores und den Leistungsdaten

variablen in der KFA	Testerr. zeitpunkt	N	signif. Typ	r	E	p	s.	p	s.
$A(u, u', u''), u(u),$ erw.Zensur, Prozentsatz, erh.Zensur	P1 - P4	140	121212 222121	17 5	4.50 1.01	.790-05 *** .920-03 *	.0351 *** .0750 ***		
$A(H), U(H), U(H),$ erw.Zensur, Prozentsatz, erh.Zensur	P1 - P4	140	121212 212111	12 7	3.25 1.43	.240-07 * .600-03 *	.0541 *** .0020 ***		
$A(u), u', u'', u(u),$ $A(H), U(H), U(H),$ erw.Zensur, Prozentsatz, erh.Zensur	P1 - P4	140	121121212	8	0.59	.180-04 ***	.0569 ***		
$A(u-H), u(u-H), u(u-H),$ erw.Zensur, Prozentsatz, erh.Zensur	P1 - P4	140	-						
$A(u), u', u'', u(u),$ erw.Zensur, Prozentsatz, erh.Zensur, Prüf.art	P1 - P4	140	1212121	18	1.34	.170-06 ***	.0910 ***		
$A(H), u', u'', u(u),$ erw.Zensur, Prozentsatz, erh.Zensur, Prüf.art	P1 - P4	140	-						
$A(u), u', u'', u(u),$ $A(H), u', u'', u(u),$ erw.Zensur, Prozentsatz, erh.Zensur, Prüf.art	P1 - P4	140	1211212121	7	0.47	.310-06 ***	.0503 ***		
$A(u-H), u(u-H), u(u-H),$ erw.Zensur, Prozentsatz, erh.Zensur, Prüf.art	P1 - P4	140	-						

Das Kollektiv derjenigen, die entspannt, variabel und analytisch in Zukunft vorgehen wollen, wird mit niedrigen Leistungsdaten bei einem Pränanzverlust von nur $D(Q) = .0058$ assoziiert. Dieses Kollektiv rekrutiert sich zu 94 % aus VDn. Die gleiche Assoziation findet sich zwischen den Leistungsdaten und den realisierten Strategien; ebenso wie zwischen Leistungsdaten, den realisierten und intendierten Strategien, wenngleich jetzt unter deutlichem Pränanzverlust gegenüber den Einzelanalysen.

Die in den Einzelanalysen erhaltenen Gegentypen, d. h. das überdurchschnittliche Leistungskollektiv und das Kollektiv der Angespant-Stetigen und Synthetisierenden bilden kein gemeinsames signifikantes Kollektiv, sondern verbinden sich mit bislang nicht-signifikanten Konfigurationen zu schwach pränanten Typen, die zudem bei Hinzufügung der Variablen Prüfungsart wieder 'verschwinden'.

Insgesamt zeigt sich also eine Teilassoziation zwischen unterdurchschnittlichen Leistungen und entspannt-variabel-analytischem Vorgehen für VDn.

Zusammenfassend fanden wir folgende Ergebnisse:

Bei der gemeinsamen Analyse von Kernstrategieitems und Leistungsdaten sind die Pränanzverluste so stark, daß eine Assoziation ausgeschlossen werden kann. Allenfalls kann man sagen, daß 'Assertive' zwar kein bestimmtes Leistungskriterium erreichen, aus der Prüfung jedoch soviel Informationen entnehmen, daß sie ihre tatsächliche Zensur gut vorhersagen können.

Die Analyse von Strategietrivialitems und Leistungsdaten ergibt, daß eindeutige Stellungnahmen zu Selbstverständlichkeiten und überdurchschnittliche Leistungsdaten einerseits, 'differenzierte' Stellungnahmen mit unterdurchschnittlichen Leistungsdaten bei VDN andererseits einhergehen.

Ein intendiertes mäßig angespanntes, variables, diversifizierendes Vorgehen geht (bei VDN) mit unterdurchschnittlichen Leistungsdaten einher.

In allen Fällen sind die Prägnanzverluste gegenüber der Einzelanalyse der Leistungsdaten zu hoch, um von den Leistungsdaten her auf Strategien schließen zu können, während die durchweg nur leichten Prägnanzverluste gegenüber den Einzelanalysen der Strategien einen Schluß von den Strategien auf die Leistungsdaten erlauben.

Laut Hypothese 4 sollen aktive und kontrollierende Strategien zum Prüfungserfolg (überdurchschnittliche Leistungsdaten) führen, während die Kandidaten mit passiven und hilflosen Strategien nicht-erfolgreich agieren. In der Hypothesentriade haben wir zudem die Kontrollstrategie in Weiterführung des Immunisierungsparadigmas in Beziehung zu überdurchschnittlichem Abschneiden und der Anthac-Klasse der HDN gesetzt und die Strategie der Gelernten Hilfllosigkeit zu den Merkmalsausprägungen 'VD' und 'unterdurchschnittliche Leistungsdaten'.

Diesbezüglich eindeutige Ergebnisse erhalten wir nur für den Satz der Trivialitems. 'Konforme' Ant-

worten auf Selbstverständlichkeiten führen zu überdurchschnittlichen Ausprägungen, 'differenzierte' zu unterdurchschnittlichen Ausprägungen der Leistungsdaten. In Weiterführung dieser Ergebnisse zeigt sich, daß auch für den Datensatz des TVD zumindest der Teil der Hypothesen, der sich auf die VDn bezieht, angenommen werden kann. Somit errechnen wir konvergente Ergebnisse, die für eine Assoziation der Merkmalsausprägung VD (der Variablen Prüfungsart), unterdurchschnittliche Leistungsdaten, mäßig angespanntes, variables, diversifizierendes Vorgehen und 'differenzierte' Stellungnahmen zu den Trivialitems sprechen.

Gleichzeitig gelingt es nicht, konvergente Ergebnisse für eine Assoziation der entgegengesetzten Merkmalsausprägungen oben genannter Datensätze zu erhalten. Wir führen dies darauf zurück, daß die entsprechenden Konfigurationen durch die Hinzufügung der Variablen Prüfungsart 'aufgespalten' werden und als Folge größtenteils die Signifikanzgrenze verfehlen.

4.3.6 Überblick und Zusammenfassung der Ergebnisse des hypothesenprüfenden Teils

Bis hierin haben wir die Ergebnisse en detail dargestellt. Im folgenden zusammenfassenden Überblick greifen wir die wichtigsten Ergebnisse noch einmal in ihrer verbalen Formulierung heraus. Als Hauptfunktion dieses Kapitels betrachten wir die Möglichkeit für den Leser, sich einen orientierenden Überblick über die Ergebnisse des hypothesenprüfenden Teils dieser Arbeit zu verschaffen.

Die Emotionstests HiS, STAI und BDI ergeben in der retrospektiven und in der On-line-Messung die gleichen KFA-Typen: Wir erhalten jeweils Kollektive Niedrig-Emotionalisierter und Hoch-Emotionalisierter.

Das gleiche Ergebnis erhalten wir in KFAn mit den beiden Analogskalen 'Erschöpfung' und 'Konzentrationsfähigkeit', die nach der Prüfung vor Bekanntgabe der Zensur von den Kandidaten bearbeitet wurden. Bei der Agglutination der Datensätze der Emotionstests, der Analogskalen und der Beschwerdeliste zeigt sich, daß sich die Träger jeweils unterdurchschnittlicher Merkmalsausprägungen und jeweils überdurchschnittlicher Merkmalsausprägungen zu Typen 'unterdurchschnittlicher Belastetheit' bzw. 'überdurchschnittlicher Belastetheit' zusammenfinden. Dies gilt auch für die reliable Fremdbeobachtervariable 'Aufregung zu Beginn', die zusammen mit den Emotionstests, nicht aber mit der Beschwerdeliste und den Analogskalen, Typen 'unterdurchschnittlicher' und 'überdurchschnittlicher' Belastung errechnen lassen. Bei der Analyse der formalen Reliabilität (von Prüfung zu Prüfung) erhalten wir Ergebnisse, die auf eine deutliche Merkmalsfluktuation (von Prüfung zu Prüfung) schließen lassen.

Bei Hinzufügung der Variablen Prüfungsart ergeben sich übereinstimmende Ergebnisse. Das Kollektiv der 'Niedrig-Emotionalisierten' wird aufgespalten und sowohl der Anthac-Klasse der VDN als auch der Anthac-Klasse der HDN zugeordnet. Das Kollektiv

der 'Hoch-Emotionalisierten' wird (bei insgesamt geringfügigen Pränanzverlusten) stets der Klasse der VDN zugeordnet.

Bei den Analysen der Testverhaltensdifferentiale, die retrospektiv wie prospektiv skaliert werden, erhalten wir häufig zwei gegenläufige Typen, die sich bei der gemeinsamen Analyse der beiden TVD-Versionen sogar zusammenfinden: einen Typ mit dem Profil 'mäßig angespannt-variabel-diversifizierend' und einen Gegentyp mit dem Profil 'angespannt-stetig-unifizierend'. In den KFAn der Kernstrategieitems erhalten wir einen Typ der 'Unsicheren' und einen Gegentyp der 'Assertiven'. Bei der Agglutination der Datensätze zeigt sich, daß die 'Assertiven' 'angespannt-stetig-unifizierend' vorgegangen sind und vorgehen wollen und daß die 'Unsicheren' 'mäßig angespannt-variabel-diversifizierend' vorgegangen sind und vorgehen wollen. Die Wiederholungsläufe unter Hinzufügung der Variablen Prüfungsart erbringen als Ergebnis, daß letzterer Typ (die 'Unsicheren', 'mäßig Angespannten-variablen-Diversifizierenden') der Klasse der VDN zuzuordnen ist, ersterer meist aufgespalten wird. Methodisch interessant ist dabei, daß die Kollektive mit o. a. Strategien einer Gruppe gegenüberstehen, die sich mit unseren Datensätzen nicht klassifizieren lassen. Ein wichtiges Ergebnis ist also, daß es Kandidaten mit Kernstrategien und Kandidaten ohne Kernstrategien gibt.

Die Trivialitems der Strategiensammlung ergeben in der KFA einen Typ 'Konformer' und einen Gegentyp, den wir 'Differenzierte' genannt haben. Die Trivialitems stehen in keinem Zusammenhang mit den anderen Strategie-Datensätzen, wenn wir davon absehen, daß 'differenzierte' und 'unsichere' Merkmalsträger überzufällig häufig zu beobachten sind.

Diejenigen Kandidaten, die eine der von uns klassifizierten Kernstrategien skalieren, läßt der Prüfer meist ausreden, was insbesondere auf die 'assertiven' VDN zutrifft. Im übrigen fanden wir zwischen den über Fremdbeobachter erhobenen 'Takiken' und den Strategien keine nennenswerten Assoziationen.

In den KFA des Datensatzes, den wir 'Leistungsdaten' genannt haben, berechneten wir ein Kollektiv 'Überdurchschnittlicher' und ein Kollektiv 'Unterdurchschnittlicher' mit den gesamtmedian-dichotomisierten Variablen 'erwartete Zensur', 'erhaltene Zensur' und 'Prozentsatz des Stoffes' (mit dem sich der Kandidat bereits während des Studiums beschäftigt hat). Im Wiederholungslauf zeigt sich, daß das Kollektiv der 'Überdurchschnittlichen' (die eine überdurchschnittliche Zensur erwarten und erhalten und einen überdurchschnittlichen 'Prozentsatz' skalieren) aufgespalten wird, während das Kollektiv der 'Unterdurchschnittlichen' der Anthac-Klasse der VDN zugeordnet wird. Die gleichen Ergebnisse erhalten wir in der Agglutinations-KFA mit den Leistungsdaten und der Zensurein-

schätzung durch den Beisitzer: In den Fällen, in denen die tatsächlich vergebene Zensur vom Kandidaten zutreffend als über- oder unterdurchschnittlich vorhergesagt wird, kommen Prüfer und Beisitzer zur gleichen Leistungsbeurteilung.

Bei den Leistungsdaten finden wir deutliche Unterschiede zwischen den Prüfern. Diese Unterschiede können auf das fehlende oder vorhandene Angebot an lernintensiven Lehrveranstaltungen in den beiden vorausgehenden Semestern zurückgeführt werden.

Kernstrategien und Belastungsmessungen stehen in keiner klaren Beziehung zueinander; wir verzeichnen deutliche Prägnanzverluste. Das Kollektiv der 'Assertiven' skaliert niedrige oder hohe Emotionen während der Prüfung, sowohl 'Assertive' wie 'Unsichere' skalieren nach der Prüfung überdurchschnittliche Erschöpfung und Konzentrationsunfähigkeit.

Realisiertes stetiges und unifizierendes Vorgehen geht mit unterdurchschnittlichen Scores, variables und diversifizierendes Vorgehen mit überdurchschnittlichen Scores simultan in HiS, STAI und BDI einher. Intendiertes angespanntes, stetiges und unifizierendes Vorgehen wird von VDN zusammen mit überdurchschnittlicher Erschöpfung und Konzentrationsunfähigkeit nach der Prüfung skaliert.

Das aus den Analysen der Strategietrivialitems bekannte Kollektiv der 'Differenzierten' (die 'Selbstverständlichkeiten' nicht eindeutig beantworten) ist auch ein Kollektiv mit überdurchschnittlichen Belastungen. Dies betrifft vor allem die Klasse der VDN.

'Assertive' erreichen zwar kein bestimmtes Leistungskriterium, entnehmen aus der Prüfung jedoch so viele Informationen, daß sie ihre tatsächliche Zensur gut vorhersagen können. Intendiertes mäßig angespanntes, variables, diversifizierendes Vorgehen geht (bei VDN) mit unterdurchschnittlichen Leistungen einher. Die Analyse von Strategietrivialitems und Leistungsdaten ergibt, daß 'Konforme' überdurchschnittliche Leistungsdaten, 'Differenzierte' aus der Klasse der VDN unterdurchschnittliche Leistungsdaten erreichen.

Überdurchschnittliche Prüfungsbelastungen, unterdurchschnittliche Leistungsdaten und die Merkmalsausprägung VD sind assoziiert. Die Ergebnisse hinsichtlich des Gegentyps mit unterdurchschnittlichen Belastungen, überdurchschnittlichen Leistungsdaten und der Merkmalsausprägung HD sind weniger eindeutig.

4.4 Hypothesengenerierender Teil

4.4.1 Analyse der einzelnen Datensätze

4.4.1.1 Die Analyse der Prüferwahlen

Laut Prüfungsordnung soll jeder Kandidat zwischen mindestens zwei Prüfern pro Fach wählen können. Zu dem von uns erfaßten Prüfungstermin konnte dieses Gebot der Prüfungsordnung vom Prüfungsamt für zwei Fächer (1 im VD, 1 im HD) nicht realisiert werden,

weil zwei Prüfungsberechtigte kurzfristig nicht zur Verfügung standen und zwei Alternativprüfer noch nicht in Frage kamen. Das bedeutete für die Kandidaten, daß die HDn und VDN in je einem Fach keine Wahlmöglichkeit hatten.

Prinzipiell bietet die Wahl eines Prüfers jedoch eine Möglichkeit, Prüfungsangst und -anforderungen zu reduzieren, zu kanalisieren oder kalkulierbar zu machen. Wir formulierten die Hypothese, daß in die Entscheidung für einen Prüfer die Kalkulation von Angst und Anforderungen einfließt, und zwar wählen die Kandidaten diejenigen Prüfer, die weniger Angst auslösen bzw. deren Anforderungen geringer sind.

Zur Überprüfung verwendeten wir dazu das Design der Abbildung 23.

Gewählt wurde	Von	Einschätzung	
		Prüfer I	Prüfer II
Prüfer I	1		
	2		
	3		
	...		
Prüfer II	...		
	...		
	...		
	...		

Abbildung 23: Analysedesign der Prüfervariablen

Zusätzlich untersuchten wir die Einschätzungen von Prüfern der Vor- und Hauptdiplomanden, die die eine Gruppe jeweils wählen konnte, die andere jedoch nicht. So konnten VDN den Prüfer G wählen, nicht jedoch Prüfer H (dieser prüfte nur im HD);

HDn konnten den Prüfer H wählen, nicht jedoch Prüfer G (dieser prüfte nur Fächer des VD).

In beiden Fällen erwarteten wir signifikante Interaktionen.

Die Analyse der Rohdaten ergab bis zu 35 % fehlende Antworten. Mögliche Vergleiche reduzierten sich also auf die Fälle, in denen ausreichend Versuchspersonen geantwortet hatten.

Analysen waren nurmehr für drei Fächer möglich. KFAn mit dichotomisierten Werten nach dem Gesamtmedian (also den mittleren insgesamt skalierten Ängsten und Anforderungen) ergaben keine signifikanten Typen. Die Hypothese, eine bestimmte Entscheidung für einen Prüfer würde mit bestimmten Skalierungen einhergehen, mußte zurückgewiesen werden. Die Hypothese, die Angst und die Prüfungsanforderungen von wählbaren und nicht wählbaren Prüfern würden unterschiedlich skaliert, wurde mit schrittweise multipler Regression geprüft. Als einziges signifikantes Ergebnis erhielten wir die signifikante Regression der Angst vor dem Prüfer G auf die Angst vor dem Prüfer H ($F=4.74$, $p < .05$). Alle Interaktionen blieben nicht-signifikant. Eine Berücksichtigung des Geschlechts und der Prüfungsart veränderte nicht die Resultate.

Sämtliche Items wurden unter Erhöhung des nominalen Alpha-Risikos auf Einflüsse der Prüfungsart und des Geschlechts über alle Kandidaten mit t-tests geprüft. Es wurden keine signifikanten Ergebnisse erhalten.

Während der Einsatz inferenzstatistischer Verfahren ergebnislos blieb, wurden wir von den korrelationsstatistischen Ergebnissen überrascht (jeweils über alle Versuchspersonen unter Berücksichtigung der Missing Values gerechnet).

Tabelle 65 gibt in der oberen Dreiecksmatrix die Korrelationen der skalierten Angst, in der unteren Dreiecksmatrix die Korrelationen der skalierten Prüfungsanforderungen und in der Diagonale die pro Prüfer errechnete Korrelation von Angst und Prüfungsanforderungen wieder.

Tabelle 65: Produkt-Moment-Korrelationen zwischen Ängsten vor Prüfern (obere Dreiecksmatrix), den skalierten Anforderungen dieser Prüfer (untere Dreiecksmatrix) und zwischen Ängsten vor und Anforderungen bei den einzelnen Prüfern (Diagonale); unterschiedliche Signifikanzniveaus ergeben sich aus unterschiedlich häufigen Missing Values (x = 5 %, xx = 1 %, xxx = 0,1 %, jeweils auf die tatsächlich verrechnete Probandenzahl bezogen)

Prüfer	A	B	C	D	E	F	G
A	.41xx	.42xx	.39xx	.56xxx	.40xx	.37x	.40x
B	.19	.57xxx	.46xx	.12	.57xxx	.11xxx	.63xxx
C	.37xx	.45xx	.42xx	.23	.61xxx	.13	.61xxx
D	.58xxx	.05	.41xx	.14xx	.00	.29	.26
E	.06	.53xxx	.66xxx	.23	.54xxx	.55xxx	.54xxx
F	.12	.22	.26x	.38x	.38x	.66xxx	.45xx
G	.32x	.51xx	.38x	.39x	.46xx	.34x	.49xx

In der Matrix finden sich nur positive Korrelationen, wenngleich einige auch nicht-signifikant ausfallen. Linkage-Analysen der Ängste, der Prüfungs-

anforderungen und der obigen Gesamtmatrix ergaben durchweg drei Cluster, also Prüfergruppen. Wichtig erscheint uns die unterschiedlich starke Assoziation zwischen Skalierung der Prüfungsanforderungen und der Angst bei den einzelnen Prüfern. Aus dem Vertrauensintervall für die höchste Korrelation (Prüfer F, $r=.66$) fällt bei Annahme eines 5%-Niveaus die niedrigste Korrelation (Prüfer D, $r=.34$) heraus, bei Annahme höherer Niveaus auch noch die Korrelation der Prüfer A und G.

Diese prüferspezifische Dissoziation bzw. Assoziation stützt unsere obigen Ergebnisse. Eine einfache Beziehung zwischen der Angst vor einem bestimmten Prüfer sowie der Einschätzung seiner Prüfungsanforderungen mit der Entscheidung für bzw. gegen einen Prüfer konnte nicht gefunden werden.

Zur Überprüfung der Mittelwertsunterschiede zwischen den Prüfern verwendeten wir t-tests. Ein varianzanalytisches Vorgehen entfiel, da bei multivariater Messung ein Missing Value zum Ausschluß der gesamten Daten einer Versuchspersonführte und somit die Verrechnungsbasis zu gering wurde.

Tabelle 66: Ergebnisse der Mittelwertsvergleiche zwischen den Prüfern. Obere Hälfte skalierte Ängste, untere Hälfte skalierte Prüfungsanforderungen. Angegeben sind die Werte der Verteilungsfunktion und welcher Prüfer bei signifikantem Unterschied den höheren Mittelwert hatte. Missing Values und ggf. ungleiche Varianzen wurden berücksichtigt.

Prüfer	A	B	C	D	E	F	G
A		.000 B	.000 C	.70	.000 E	.000 F	.06
B	.000 B		.69	.000 B	.16	.07	.000 B
C	.000 C	.001 B		.000 C	.04 C	.14	.000 F
D	.01 A	.000 B	.000 C		.000 E	.000 F	.000 F
E	.002 E	.000 B	.78	.000 E		.74	.001 E
F	.04 F	.000 B	.07	.000 F	.22		.025 F
G	.68	.000 B	.000 C	.13	.000 E	.001 F	

Obwohl aus Tabelle 66 eindeutig hervorgeht, daß die Prüfer unterschiedlich eingeschätzt wurden, gelang es nicht, damit einhergehende Wahlen aufzuzeigen. Zu vermuten ist, daß in die Einschätzung der Prüfer und in die Prüferwahl weitere, von uns nicht erfaßte Variablen eingehen.

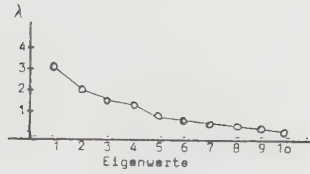
4.4.1.2 Die Attribuierungsskalen

In der BL-Erhebung hatten wir die zwölf Attribuierungssitems mit zwei verschiedenen Instruktionen eingesetzt: "Gesetzt den Fall, Du würdest besser als erwartet abschneiden, worauf würdest Du die Ursachen zurückführen?" sowie in der analogen Formulierung für schlechter als erwartetes Abschneiden. Nach den Prüfungen ließen wir das tatsächliche Abschneiden beurteilen. Die vorgesehene Kennwertbildung überprüften wir faktorenanalytisch. In Abbildung 24 sind die Eigenwertverläufe für zwei Faktorenanalysen mit je 12 Items pro Instruktion ('besser' und 'schlechter') und einer Faktorenanalyse mit allen 24 Items (verkleinert) abgebildet ('besser und schlechter').

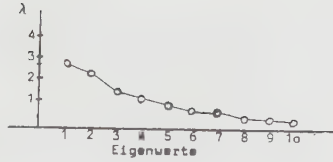
Nach dem Scree-Test hätten wir bei der 'schlechten' Instruktion einmal 4 (bzw. 6) Faktoren, bei der 'besser als erwartet'-Instruktion 2 (bzw. 4) Faktoren, nach dem Verlauf der 24er Lösung 5 Faktoren extrahieren müssen.

Wir entschlossen uns daher, die Faktorenanalyse mit 140 Daten der Datenerhebungszeitpunkte P1 bis P4 zu wiederholen, dabei Meßwiederholungen nicht zu berücksichtigen.

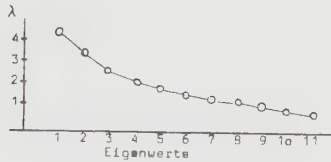
'schlechter'



'besser'



'schlechter'
und
'besser'



N = 140

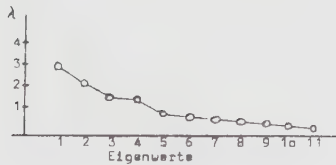


Abbildung 24: Eigenwertverlauf aus Hauptkomponentenanalysen der Attribuierungsskalen; links neben den Darstellungen Kurzbezeichnungen für die verrechneten Datensätze; Näheres im Text

In Abbildung 24 ist der Eigenwertverlauf dieser Faktorenanalyse und in Tabelle 67 die varimaxrotierte Lösung dargestellt.

Tabelle 67: Matrix einer varimaxrotierten Hauptkomponentenlösung der Attribuierungsitems mit vier extrahierten Faktoren bei $N = 140$

Item	Text	Faktor 1	Faktor 2	Faktor 3	Faktor 4
IV1+	Ich habe mich angestrengt	-.46			
IV2-	Ich war faul	.74			
IV3+	Ich war fleißig	-.77			
EV1+	Ich hatte Glück				-.82
EV2-	Ich hatte Pech				.73
EV3+	Das war Zufall			-.64	(-.34)
IK1+	Ich bin ein glücklicher Mensch			.23	
IK2-	Das Leben ist voll von Schwierigkeiten			-.31	.43
IK3+	Ich bin intelligent			.73	
EK1+	Der Prüfer prüft gut		.90		
EK2-	Der Prüfer prüft schlecht		-.91		
EK3+	Es wird nicht viel verlangt	.65	(.02)		

Sowohl Scree-Test als auch Kaiserkriterium ergeben für die Faktorenanalyse eine vierfaktorielle Lösung. Nach der Varimaxrotation finden sich vier Faktoren (es wurden nur Faktorladungen $a(i) > |.40|$ berücksichtigt).

Faktor 1: "Ich habe mich angestrengt" (bestimmt nicht) - "Ich war faul" (völlig sicher) - "Ich war fleißig" (bestimmt nicht) - "Es wird nicht viel verlangt" (völlig sicher);
internal-variable Attribuierung.

Faktor 2: "Der Prüfer prüft gut" (völlig sicher) - "Der Prüfer prüft schlecht" (bestimmt nicht);
external-konstante Attribuierung.

Faktor 3: "Das war Zufall" (bestimmt nicht) - "Es liegt an meinen Fähigkeiten" (völlig sicher) - "Ich bin intelligent" (völlig sicher) - ("Das liegt an meiner Unfähigkeit" (vermutlich nicht);
internal-konstante Attribuierung.

Faktor 4: "Ich hatte Glück" (bestimmt nicht) - "Ich hatte Pech" (völlig sicher) - "Das liegt an meiner Unfähigkeit" (völlig sicher) - ("Das war Zufall" (vermutlich nicht);
external-variable Attribuierung.

In Tabelle 68 ist die der Hauptkomponentenanalyse zugrunde gelegte Korrelationsmatrix als untere Dreiecksmatrix wiedergegeben.

Tabelle 68: Multi-Trait-Multi-Method-Matrix der Attribuerungsitems für N = 140 (P1 bis P4)

Items	IV1+	IV2-	IV3+	EV1+	EV2-	EV3+	IK1+	IK2-	IK3+	EK1+	EK2-	EK3+
IV1+												
IV2-	-.26											
IV3+	.25	-.54										
EV1+	-.07	.07	-.23									
EV2-	-.16	.26	-.01	-.41								
EV3+	-.25	.21	-.14	.39	.09							
IK1+	.31	-.15	.08	-.04	-.19	-.41						
IK2-	-.27	.34	-.14	-.10	.32	.20	-.22					
IK3+	.20	-.17	.14	-.12	-.09	-.21	.41	-.20				
EK1+	.19	-.02	-.01	.20	-.25	-.13	.20	-.18	.02			
EK2-	-.14	.10	.01	-.21	.29	.07	-.13	.21	.06	-.75		
EK3+	-.27	.20	-.27	.25	.04	.10	.05	.12	-.06	.09	-.11	

Wenn wir diese Matrix als Multi-Trait-Multi-Method-Matrix auffassen, so zeigt sich, daß keine der Heterotrait- oder Heteromethodkorrelationen absolut größer als die größte Intratraitkorrelation ist. Ansatzweise können wir demnach von konvergierender Validität der Items sprechen.

Die vierfaktorielle Lösung der Faktorenanalyse deutet auf vier unabhängig variierende Attribuerungsaspekte. Für die weiteren Analysen müssen wir daher berücksichtigen, daß die Versuchspersonen möglicherweise auf mehreren Dimensionen gleichzeitig attribuieren.

Beispiel: "Das liegt an meiner Fähigkeit" und "Ich hatte Glück"; "Ich habe mich angestrengt" und "Der Prüfer prüft gut". Dadurch wird es mög-

lich, daß das lineare Modell der Faktorenanalyse verletzt wird, z. B. internal-variable Attribuierung zur Voraussetzung der internal-konstanten wird ("Ich bin intelligent, weil ich mich (schon lange vorher) angestrengt habe"). Ferner ist es möglich, bestimmte Attribuierungsmöglichkeiten auszuschließen ("Das war kein Zufall, aber liegt es am Prüfer oder an mir ... ?")

Wir verzichteten deshalb auf die Berechnung von Faktorscores und bildeten vier Attribuierungsscores; bei Ablehnung eines Items vergaben wir wachsend negative Werte, bei Zustimmung positive; wir addierten die beiden positiv formulierten Items und subtrahierten das negativ formulierte.

In Tabelle 69 sind die erhaltenen Daten tabelliert.

Die Eigenart unserer Skalenkonstruktion läßt es zu, daß die Versuchspersonen einem Attribuierungsaspekt zustimmen oder ihn ablehnen. Die Ergebnisse der BL-Datenerhebung zeigten, daß die Versuchspersonen diese Möglichkeit ausnutzten: die Kausalattribution für ein 'besser als erwartetes' Abschneiden geht in Richtung Fähigkeit, gutes Prüfungssystem, Anstrengung und Glück, die Kausalattribution für ein 'schlechter als erwartetes' Abschneiden geht in Richtung Unfähigkeit, schlechtes Prüfungssystem, Faulheit und Pech. Die Unterschiede zwischen den beiden Instruktionsbedingungen sind jeweils auf dem 1-%-Niveau signifikant.

Interessant ist nun, daß die Kausalattribution der tatsächlich erhaltenen Zensuren der der fiktiv-

Tabelle 69: Mittelwerte und Standardabweichungen der Attribuierungsskalen

Dimension	Datenerh.zeitpunkt	M	s	n(x) < 0; in %
IK	BL ("besser")	3,78	1,80	4 9,8%
	BL ("schlechter")	-0,22	2,22	24 58,5%
	P1	3,66	2,03	1
	P2	3,45	2,42	5
	P3	3,44	2,16	3
	P4	3,53	2,39	2
	P1 - P4 (N=140)	3,51	2,21	11 7,8%
EK	BL ("besser")	3,17	2,34	6 14,6%
	BL ("schlechter")	-2,63	2,08	41 100,0%
	P1	2,73	2,66	6
	P2	2,29	3,00	10
	P3	2,15	2,75	11
	P4	-0,19	3,95	10
	P1 - P4 (N=140)	1,76	3,17	37 26,4%
IV	BL ("besser")	4,43	2,09	3 7,3%
	BL ("schlechter")	-1,00	3,50	28 68,7%
	P1	4,01	2,96	4
	P2	4,24	2,83	5
	P3	4,61	2,44	1
	P4	3,68	2,50	2
	P1 - P4 (N=140)	4,32	2,71	14 10,0%
EV	BL ("besser")	3,90	2,10	3 7,3%
	BL ("schlechter")	-2,29	1,19	41 100,0%
	P1	3,78	3,27	15
	P2	2,17	2,99	12
	P3	1,07	3,21	16
	P4	1,37	3,10	7
	P1 - P4 (N=140)	1,11	3,13	50 35,1%

ven Erfolgsattribuierung ('besser als erwartet')
- besonders in ~~den~~ internalen Dimensionen - ähnelt.

Die multivariaten Zusammenhänge überprüften wir wieder konfigurationsfrequenzanalytisch. Mit den Variablen IK, EK, IV, EV und Prüfungsart rechneten wir für die Gesamtdaten und jeden Datenerhebungszeitpunkt gesondert alle Zweier-, Dreier-, Vierer- und Fünfer-KFAn, die möglich sind. Insgesamt erhielten wir in 130 Läufen ganze drei signifikante Typen und zwar: für alle Daten unter der Annahme der unabhängigen Messung für die Variablen IK, EK, IV, EV ~~den Typ~~ 2222 mit $F=19$, $E=8.97$, $B=.11D-O2(x)$ und $Q=.0787(xxx)$ - also ein Kollektiv von Merkmalsträgern, das sowohl internal-konstant als auch external-konstant als auch external-variabel als auch external-variabel attribuiert -, für den Datenerhebungszeitpunkt P1 mit $N = 41$ und für die Variablen IK, IV, EV, Prüfungsart den Typ 1112 mit $F=5$, $E=1.04$, $B=.32D-O2(-)$ und $Q=.0991 (xxx)$ - also ein auf dem 10-%-Niveau signifikantes zu vermutendes Kollektiv von Hauptdiplomanden, die die IK, IV und EV-Attribuierung zurückweisen. Bei der Analyse der BL-Daten erhielten wir nur einen signifikanten Typ. In einem Lauf mit sämtlichen acht Attribuierungsskalen (vier Dimensionen und zwei Instruktionen) stießen wir auf ein Kollektiv, das im Erfolgsfall auf beiden externalen Dimensionen positiv, im Mißerfolgsfall nur external-variabel positiv attribuiert (Konfiguration: 12121112, $F=4$, $E=0.23$, $B=.91D-O4x$, $Q=.0924xxx$), mithin auf ein Kollektiv, das bei besser als erwartetem Abschneiden dies auf die Institution und auf Glück, nicht auf Fähigkeit und Anstrengung zurückführt,

die Ursache für schlechter als erwartetes Abschneiden allein im Glück, nicht in Fähigkeit, Anstrengung oder der Situation sieht.

Tabelle 70: Ergebnisse der KFAn mit den Attribuierungsskalen

Attribution in neg. KFA	N	signif. Vc	F	S	B	S.	Z S.
IK,EK,IV,ZV	140	100%	19	3,60	.110-02	x	.0787 xxx
IK,EK,IV,ZV, erh.Zensur	140	100%	19	3,60	.140-07	xxx	.1040 xxx
IK,EK,IV,ZV, erh.Zensur, Datenerh.ztkt.,Prüf.art	140	100% 100%	4 5	2,00 2,37	.130-03 .130-05	x xxx	.0507 xxx .0505 xxx
IK,EK,IV,ZV, erh.Zensur, Datenerh.ztkt.,Prüf.art	140	100% 100%	7 6	0,55 1,39	.160-04 .150-03	xx x	.0476 xxx .0357 xxx
IK,EK,IV,ZV, erh.Zensur, Datenerh.ztkt.,Prüf.art	140	-					

Die Analyse der Attribuierungsverläufe - also der formalen Reliabilitäten - ergab sowohl bei Verrechnung der Differenzwert- als auch der Rohdaten (gesamtmediandichotomisiert) keinen signifikanten Typ. Der oben erwähnte signifikante Typ entsteht also nicht durch mehrfache Verrechnung eines hochreliabel attribuierenden Minikollektivs.

Die weitere Analyse ist Tabelle 70 zu entnehmen. Danach besteht das Kollektiv der positiv Attribuierenden zu 89 % aus Versuchspersonen, die überdurchschnittlich abschnitten (entsprechend unserer Medianisierung also eine '1' bekamen); dieses

Teilkollektiv erhielt zu 76 % (insgesamt 68 %) seine Zensur in den beiden ersten Prüfungsterminen. Jeweils rund die Hälfte aller '1'er-Kandidaten in den Datenerhebungszeitpunkten attribuierte auf allen Dimensionen positiv. Dieses Attribuierungsmuster verliert sich im weiteren Verlauf der Prüfungen.

4.4.1.3 Die emotionale Reaktion auf die Zensur

Unmittelbar nachdem die Kandidaten ihre Zensur erfahren hatten, wurde ihnen der Fragebogen F2 vorgelegt. Eine der Fragen war: "Welche Gefühle löst die Zensur in Dir aus? (Bitte alles Zutreffende ankreuzen!)" Es wurden zwölf Antwortmöglichkeiten vorgegeben, von denen nach Vorinterviews des Autors angenommen werden konnte, daß sie das gesamte Spektrum der unmittelbaren Emotionen (= Konsequenzen) abdeckten. In der folgenden Tabelle sind die absoluten und relativen Häufigkeiten, getrennt nach Datenerhebungszeitpunkten, wiedergegeben.

Tabelle 71: Absolute und relative Häufigkeiten der durch die Zensur ausgelösten Emotionen

Emotion	N(VP)					%(VP)				
	P1	P2	P3	P4	P1-P4	P1	P2	P3	P4	P1-P4
1. Erleichterung	30	30	32	16	108	73	73	82	84	77
2. Freude	26	28	24	12	90	63	68	62	63	64
3. Ärger	5	6	4	3	18	12	15	10	16	13
4. Enttäuschung	7	7	7	4	25	17	17	18	21	18
5. Wut	2	5	2	2	11	5	12	5	11	8
6. Angst	0	0	0	0	0	-	-	-	-	-
7. Depress. Gef.	1	1	1	1	4	2	2	3	5	3
8. Zufriedenheit	20	21	22	10	73	49	51	56	53	52
9. Zuversicht	17	16	17	7	57	41	39	44	37	41
10. Sicherheit	14	10	17	7	48	34	24	44	37	34
11. Verwirrung	2	7	1	2	12	5	17	3	11	8
12. Hilflosigkeit	1	1	2	0	4	2	2	5	-	3

Unter der Annahme der unabhängigen Messungen nahmen wir die Menge der Ankreuzungen über alle Prüfungen pro Item und berechneten damit den Durchschnitt über alle Items. Diesen Mittelwert nahmen wir als Schätzung der Antworthäufigkeit pro Item und testeten die Ankreuzungssummen über alle Prüfungen gegen die Nullhypothese mit dem Chi-Quadrat-test und $df=1$. Danach wurden die Emotionen Erleichterung, Freude und Zufriedenheit überzufällig häufig angekreuzt, die Emotionen Ärger, Enttäuschung, Wut, Angst, Depressive Gefühle, Verwirrung und Hilflosigkeit überzufällig selten (jeweils auf dem 0,1-%-Niveau).

Sieht man von dem Item Erleichterung ab, dann werden die Zensuremationen selten durch negative Emotionen getönt, im Gegenteil: Freude und Zufriedenheit herrschen vor. Die routinemäßige Überprüfung auf Unterschiede zwischen VDN und HDN ergab lediglich für den Datenerhebungszeitpunkt P3 für das Item 'Erleichterung' einen signifikanten Wert.

Die multivariate Analyse bereiteten wir wieder mit einer Linkage-Analyse vor, wobei wir dieses Mal den Chi-Quadratwert als Verkürzung des Kontingenzkoeffizienten als Assoziationsmaß zugrunde legten.

Wir erhielten ein Cluster mit den Items Erleichterung, Freude, Zuversicht, Zufriedenheit und Sicherheit sowie ein zweites Cluster mit den Items Ärger, Enttäuschung und Wut. Auf die Darstellung verzichteten wir, da sie noch konfigurationsfrequenzanaly-

tisch untersucht werden.

In die KFAn gingen die Items Hilflosigkeit, Angst und Depressive Gefühle nicht ein, da die Fallzahlen zu gering waren. Durch Inspektion der Rohdaten zeigte sich, daß letztere drei Items nicht wiederholt von der gleichen Versuchsperson angekreuzt wurden.

In Tabelle 72 sind die Ergebnisse der KFAn wiedergegeben. In der oberen Hälfte der Tabelle sind Ergebnisse aus Läufen mit den neun verbliebenen Emotionen angegeben. Da für die weitere Verrechnung eine Reduktion der Variablenzahl aus technischen Gründen notwendig war, sind zwei weitere Läufe mit reduzierten Variablenzahlen dargestellt. Die vollständige und die reduzierte Variablenkombination wiederholten wir unter Hinzufügung der Variablen Prüfungsart; diese Ergebnisse sind im unteren Teil dargestellt. Bei der Darstellung der signifikanten Typen bedeutet '0' 'nicht angekreuzt' und '1' 'angekreuzt'.

Bei der multivariaten Analyse aller neun Emotionen über alle Datenerhebungszeitpunkte finden wir eine stark und drei schwach prägnante Überzufällige Konfigurationsfrequenzen. Wir finden einen Typ, der nur 'enttäuscht' ist, einen zweiten Typ, der nur 'Enttäuschung' und 'Ärger' ankreuzt sowie einen dritten Typ, der nur die Emotionen 'Enttäuschung' und 'Ärger' und 'Wut' als gegeben skaliert. Der vierte Typ kann als Gegenteil zu den anderen dreien aufgefaßt werden; er kreuzt die Emotionen 'Erleichterung', 'Freude', 'Zufriedenheit', 'Zuversicht'

Tabelle 72: Ergebnisse der KFA mit den Items der emotionalen Reaktion auf die Zensur

Variablen in der KFA	Datenerh. zeitpkt.	signif. Typ	F	E	B	S.	C	S.
1/2/3/4/5/8/9/10/11	P1 - P4	000100000 001100000 001110000 110001110	7 3 3 26	0,31 0,05 0,00 2,10	,310-07 ,140-04 ,300-21 ,130-16	xxx xx xxx xxx	,0474 ,0211 ,0643 ,1589	xxx xxx xxx xxx
	P1	000100000 001110000 001110000	4 2 4	0,11 0,001 0,002	,450-35 ,290-06 ,100-12	xx xxx xxx	,0852 ,0486 ,0975	xxx xxx xxx
	P2	110001110 001100000	7 7	0,11 0,06	,760-01 ,120-04	xxx x	,1502 ,0755	xxx xxx
	P3	001110000 001110000 110001110	2 2 1	0,00 1,16	,670-07 ,250-07	xxx xxx	,0613 ,2591	xxx xxx
	P4	-	-	-	-	-	-	-
1/2/3/4/ 9/9/10	P1 - P4	0001000 0011000 1100111	7 13 1	0,10 0,0 1,30	,100-06 ,010-24 ,700-15	xxx xxx xxx	,047 ,082 ,1557	xxx xxx xxx
	P1	0001000 0011000 1100111	4 2 4	0,1 0,00 0,00	,650-34 ,140-01 ,160-07	xxx x x	,0540 ,0003 ,1050	xxx xxx xxx
	P2	0011000 1100111 0011000	5 7 1	0,00 0,7 0,11	,100-05 ,110-04 ,100-01	xxx xxx xxx	,1501 ,0755 ,0767	xxx xxx xxx
	P3	0011000 1100111 0011000	1 1 1	0,00 1,00 1,00	,100-01 ,100-01 ,100-01	xxx xxx xx	,0767 ,0767 ,0767	xxx xxx xxx
	P4	0011000	1	0,00	,100-01	xx	,0767	xxx
1/2/3/4/ 9/10	P1 - P4	0001000 0011000 1100111	7 13 1	0,1 0,0 1,30	,100-06 ,010-24 ,700-15	xxx xxx xxx	,047 ,082 ,1557	xxx xxx xxx
	P1	0001000 0011000 1100111	4 2 4	0,1 0,00 0,00	,650-34 ,140-01 ,160-07	xxx x x	,0540 ,0003 ,1050	xxx xxx xxx
	P2	0011000 1100111 0011000	5 7 1	0,00 0,7 0,11	,100-05 ,110-04 ,100-01	xxx xxx xxx	,1501 ,0755 ,0767	xxx xxx xxx
	P3	0011000 1100111 0011000	1 1 1	0,00 1,00 1,00	,100-01 ,100-01 ,100-01	xxx xxx xx	,0767 ,0767 ,0767	xxx xxx xxx
	P4	0011000	1	0,00	,100-01	xx	,0767	xxx
1/2/3/4/ 9/10/ Prüf.zeit	P1 - P4	0001000 0011000 0011100002 1100011101 1100011102	2 2 16 16 10	0,1 0,1 0,0 1,0 0,1	,100-06 ,100-06 ,490-06 ,140-09 ,360-07	xxx xxx xxx xxx xxx	,047 ,047 ,047 ,100 ,047	xxx xxx xxx xxx xxx
	P1	0001000 0011000 0011100001	1 1 1	0,0 0,0 0,0	,650-34 ,430-09 ,200-05	- xxx xx	,0714 ,0714 ,0714	xxx xxx xxx
	P2	0011000 1100111 1100112	5 8 11	0,1 0,0 0,1	,170-07 ,490-16 ,350-00 ,100-04	xxx xxx xxx xx	,047 ,047 ,047 ,047	xxx xxx xxx xxx
	P3	0001001 0011000 0011002	1 1 1	0,1 0,0 0,0	,610-03 ,350-05 ,250-06	- xxx xxx	,064 ,071 ,071	xxx xxx xxx
	P4	0011002 1100111 0011002	1 1 1	0,1 1,0 0,1	,120-04 ,340-03 ,620-04	xx x xx	,047 ,064 ,1047	xxx xxx xxx

und 'Sicherheit' an. Die Emotion 'Verwirrung' gehört nicht zu den Typen.

Bei der Analyse innerhalb der einzelnen Datenerhebungszeitpunkte zeigt sich, daß die 'Nur-Enttäuschten' sich aus den Datenerhebungszeitpunkten P1 und P3 rekrutieren, was zu einem Anstieg der Prägnanzkoeffizienten führt. Die Enttäuscht-verärgert-Wütenden rekrutieren sich aus allen vier Datenerhebungszeitpunkten, sind dabei beim Datenerhebungszeitpunkt P2 häufiger als im Schnitt.

Die 'angenehm Berührten' sind in den Datenerhebungszeitpunkten P2 und P3 überrepräsentiert.

Die Hinzufügung der Variablen Prüfungsart (VD=1/HD=2) läßt die Typen zerfallen. Lediglich die 'Nur-Enttäuschten' repräsentieren ein Kollektiv von Hauptdiplomanden. Sieht man von letzteren einmal ab, so kann von den Zensuremotionen eindeutig nicht auf eine der Prüflingsgruppen geschlossen werden.

Interessant ist, daß in den dargestellten Daten eindeutig sich eine Emotionshierarchie zeigt. Das Kollektiv derjenigen, die Enttäuschung ankreuzten, zerfällt in ein Subkollektiv derjenigen, die nur 'Enttäuschung' skalierten, und ein zweites Subkollektiv derjenigen, die 'Enttäuschung' und 'Ärger' empfanden. Das Subkollektiv der 'Enttäuscht-Verärgerten' wiederum zerfällt in ein Sub-Subkollektiv von nur 'Enttäuscht-Verärgerten' und eine Gruppe von Versuchspersonen, die die Items 'Enttäuschung', 'Ärger' und 'Wut' markierten. Wir können daraus schließen,

daß 'Ärger' 'Enttäuschung' impliziert und das Item 'Wut' sogar 'Ärger' und 'Enttäuschung' impliziert. Anders formuliert: 'Enttäuschung' ist Voraussetzung für 'Ärger', und die Kombination 'Enttäuschung' und 'Ärger' ist Voraussetzung für 'Wut'.

In Abbildung 25 haben wir diese Verhältnisse versucht darzustellen.

Item	Antworten bei Kollektiv			
	I	II	III	IV
Enttäuschung	nein	ja	ja	ja
Ärger	nein	nein	ja	ja
Wut	nein	nein	nein	ja

Abbildung 25: Emotionale Hierarchie in der emotionalen Reaktion auf die Zensur

Ähnliche hierarchische, wenngleich komplexere Verhältnisse fanden wir beim Gegentypen; Zufriedenheit kann als Voraussetzung für Zuversicht und Sicherheit angesehen werden.

Im Rahmen dieser Arbeit verwendeten wir die hierarchischen Verhältnisse lediglich, um die Zahl der Variablen zu reduzieren. Wir nahmen an, daß eine Variable, die zu zwei Dritteln eine Implikation einer anderen darstellt, redundant ist.

In Tabelle 72 sind die Ergebnisse der Reduktion auf sechs Variablen wiedergegeben. Es zeigt sich,

daß die Prägnanzkoeffizienten für die Gruppen der Nur-Enttäuschten und der angenehm Berührten nahezu gleich bleiben, d. h. der reduzierte Variablenpool ist praktisch gleich brauchbar. Bei der Kombination der beiden Sub-Subkollektive zum Kollektiv der Enttäuscht-Verärgerten steigt die Prägnanz.

4.4.1.4 Die Prüfungsanmeldemotivation

Mit zwei Items hatten wir in der BL-Erhebung versucht, die Prüfungsanmeldemotivation zu erfassen. Die Items lauteten:

- "Warum hast Du Dich für die Prüfungen gemeldet?"
und
- "Gesetzt den Fall, Du müßtest eine Prüfung wiederholen, was würde geschehen?"

Jedem Item waren von uns zehn Antwortmöglichkeiten zugeordnet. In Tabelle 73 sind die erhaltenen Häufigkeiten wiedergegeben.

Die mittlere Häufigkeit verwendeten wir als theoretisch erwartete Häufigkeit und berechneten so mit dem Chi-Quadrat-Test mit $df=1$ überzufällig häufige bzw. seltene Antworten. Insgesamt werden die Items 1, 6, 16 und 18 überzufällig selten, die Items 3, 4 und 10 überzufällig häufig angekreuzt. Signifikante Unterschiede zwischen Vor- und Hauptdiplomanden fanden sich bei den Items 2, 3, 4 und 12.

Tabelle 73: Häufigkeiten und Gruppenunterschiede in den Angaben zu den Prüfungsüberzielen

Variable (Text)	N(VD)	N(HD)	Chi ²
1. weil man üblicherweise genau in meinem Semester die Prüfungen ablegt,	4	0	
2. weil ich in den Beruf will,	3	9	x
3. aus finanziellen Gründen,	10	9	x
4. damit ich mit meiner Ausbildung weitermachen kann	27	1	xxx
5. weil ich sowieso schon zu lange studiert habe,	9	4	
6. weil man mir eine Stelle angeboten hat	0	1	
7. weil das bisherige Studium langweilig geworden ist	8	2	
8. weil meine Angehörigen (..) das erwarten	9	3	
9. weil ich ausreichend vorbereitet bin	8	1	
10. weil ich etwas anderes machen möchte	10	8	
11. Meine Angehörigen waren ausgesprochen ärgerlich.	5	1	
12. Ich käme in finanzielle Schwierigkeiten	8	1	v
13. Ich wäre vor lauter Scham sehr unsicher im Umgang mit anderen.	7	1	
14. Das würde wegen des erhöhten Arbeitsaufwandes für mich die Grenze der Belastbarkeit darstellen.	5	3	
15. Das würde für mich <u>wirklich</u> nichts anderes bedeuten als ein verplempertes Semester.	8	2	
16. Ich würde zwar die Wiederholungsprüfung anstreben, wäre aber unsicher, ob ein Weiterstudium noch Zweck hat.	4	0	
17. Im ersten Augenblick würde ich die Welt nicht mehr verstehen; darüber würde ich aber hinwegkommen.	9	1	
18. Das ist eine unbeantwortbare Frage, weil die Wahrscheinlichkeit durchzufallen für mich nur theoretisch besteht.	1	2	
19. Das würde mich nicht wundern, weil ich für einige Prüfungen sowieso ein Risiko eingekalkuliert habe.	11	2	
20. Das wäre sehr ärgerlich, würde mich aber nicht umwerfen, weil ich es ja eigentlich nicht verhindern konnte.	10	5	

Auf der Basis der Chi-Quadratwerte, die wir durch Kreuztabellierung der zwanzig Items und der Variablen Prüfungsart aller gegen alle erhielten, führten wir eine Linkage Analyse durch. Eines der gefundenen Cluster ließ sich mit der KFA nicht wiederfinden: das Zweier-Cluster der Items 9 und 20. Die KFA für die beiden anderen Cluster sind in Tabelle 74 wiedergegeben.

Tabelle 74: Ergebnisse der KFA mit den Angaben zu den Oberzielen der Prüfungsmotivation

Variablen in der KFA	signif. Typ	F	L	B	S.	S.
8/13/14/17	11 1	2	0.11	.567=02	-	.0152 xxx
8/13/14/17/Prüf.art	11111	2	0.03	.296=02	-	.005 xxx
9/13/14/ Prüf.art	-					
2/3/4/10/12/Prüf.art	001001 110*12	5 4	2.13 3.13	.110=03 .060=05	x xxx	.1708 xxx .0800 xxx
2/3/4/ 12/Prüf.art	00101 11012	13 4	7.72 0.28	.110=03 .190=03	xx xx	.1105 xxx .1212 xxx
3/4/ 12/Prüf.art	0101 1012	15 5	0.11 2.04	.020=01 .020=01	xx xx	.1105 xxx .1105 xxx
6/13/14/3/4/12/Prüf.art	0000101 0001012	11 4	2.13 0.11	.110=03 .110=03	x x	.1105 xxx .1105 xxx
8/13/14/3/4/12	-					

Das Cluster der Variablen 8, 13, 14 und 17 konnte lediglich auf dem 10-%-Niveau gesichert werden. Zudem ist der Prägnanzkoeffizient ziemlich klein. Wir werten dies deshalb als Hinweis auf ein Kollektiv von Versuchspersonen, die auf ein Durchfallen in der Prüfung mit Unsicherheit und Scham reagieren würden und aus sozialen Gründen in die Prüfung gehen.

Das andere Cluster bestand aus den Kernclustervariablen

riablen 3, 4, 12 und Prüfungsart, denen die Variablen 2 und 10 als Kette zugeordnet wurden. Die Vierer-, Fünfer- und Sechser-KFAn sind in Tabelle 74 dargestellt. Das Kerncluster schlägt sich in zwei prägnanten Kollektiven nieder: ein Kollektiv von Hauptdiplomanden, das aus finanziellen Gründen in die Prüfung geht, bei einem evtl. Durchfallen in finanzielle Schwierigkeiten käme, dabei nicht angibt, mit der Ausbildung weitermachen zu wollen; und ein Kollektiv von Vordiplomanden, das nicht aus finanziellen Gründen in die Prüfung geht und auch nicht im Falle eines Falles in finanzielle Schwierigkeiten käme, sondern sich zur Prüfung gemeldet hat, um mit der Ausbildung weitermachen zu können. Fügt man Variable 2 ("weil ich in den Beruf will") hinzu, so sinken beide Prägnanzkoeffizienten, d. h. daß diese Variable zur Aussagekraft nichts hinzufügt. Fügt man nun noch Variable 10 hinzu, so sinkt der Prägnanzkoeffizient für die VDN weiter, für die HDN bleibt er in etwa gleich. Daraus kann gefolgert werden, daß für die HDN diese Variable das gleiche aussagt wie Variable 2, oder anders ausgedrückt: Die Variablen 2 ("weil ich in den Beruf will") und 10 ("weil ich etwas anderes machen möchte") sind für HDN stark assoziiert.

Aus motivtheoretischer Sicht haben wir also zwei Kollektive mit gegentypischen Oberzielen vor uns! Die Kombination des zweiten mit dem dritten Cluster macht dies noch deutlicher: Die VDN gehen einfach deshalb in die Prüfung, um mit der Ausbildung weiter zu machen (keine andere Konfigurationsfrequenz fällt in das Vertrauensintervall dieses Typs), die HDN gehen vorwiegend aus ökonomischen Motiven in

die Prüfung (eine eindeutige Aussage ist wegen der niedrigen Konfigurationsfrequenz und des damit zusammenhängenden Konfidenzintervalles nicht möglich).

4.4.2 Agglutination ausgewählter Datensätze

Im folgenden wollen wir die Ergebnisse von Agglutinations-KFAn der Attribuierungsdaten, der emotionalen Reaktion auf die Zensur und der Leistungsdaten darstellen.

In Tabelle 75 haben wir die Ergebnisse der gemeinsamen Analyse der drei Leistungsvariablen und der emotionalen Einschätzung der Zensur wiedergegeben.

Wieder haben wir ein Ankreuzen in der Emotionsliste mit '1', ein Nichtankreuzen mit '0' dargestellt. In der oberen Hälfte haben wir die Ergebnisse für den teils reduzierten und den ganz reduzierten Variablensatz der Emotionen wiedergegeben. Wir beziehen uns zunächst auf den völlig reduzierten Variablensatz.

Gegenüber beiden Einzelanalysen zeigt sich ein deutlicher Abfall der Prägnanzkoeffizienten mit Ausnahme der Analysen für den Datenerhebungszeitpunkt P4. Im allgemeinen können wir also nicht von einem Prägnanzgewinn reden.

Gegenüber der Einzelanalyse der Leistungsdaten fällt zunächst auf, daß wir signifikante Ergebnisse in den Datenerhebungszeitpunkten P2 und P4 erhalten, daß andererseits die signifikanten Ergeb-

Tabelle 75: Ergebnisse der KFA mit den Leistungsdaten und den Items der emotionalen Reaktion auf die Zensur

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. Typ	F	E	B	s.	d	s.
1,2,3,4,8,9,10, erw.Zensur, Prozentsatz, erh.Zensur	P1 - P4	131	0011000112 0011000212 0011000222 1100111121 1100111111	5 4 2 8 7	0.01 0.01 0.01 0.17 0.09	.670-12 .880-09 .765-04 .120-10 .210-07	xxx xxx x xxx xxx	.0361 xxx .0304 xxx .0152 xxx .0598 xxx .0514 xxx	
	P1	40	-						
	P2	33	0011000212	3	0.01	.670-07	xxx	.0907 xxx	
	P3	39	0011000112 1100111121	2 6	0.001 0.17	.750-05 .100-07	xxx xxx	.0513 xxx .1506 xxx	
	P4	19	0011000112	2	0.002	.240-05	xx	.1052 xxx	
1,2,3,4,9,10, erw.Zensur, Prozentsatz, erh.Zensur	P1 - P4	131	1100111111 0011001112 1100111121 001100212 001100222	7 5 8 4 2	0.05 0.02 0.37 0.01 0.01	.200-05 .140-10 .540-08 .100-07 .110-03	xx xxx xxx xxx -	.0476 xxx .0368 xxx .0564 xxx .0294 xxx .0147 xxx	
	P1	40	-						
	P2	33	001100212	3	0.01	.590-06	xxx	.0510 xxx	
	P3	39	001100112 1100111121	2 6	0.001 0.28	.350-05 .370-06	xx xxx	.0510 xxx .1467 xxx	
	P4	19	001100112	2	0.005	.170-04	xx	.1070 xxx	
1,2,3,4,8,9,10, erw.Zensur, erh.Zensur, Prüf.art	P1 - P4	131	0011000221 0011000122 0011000222 1100111111 1100111112 1000000221	4 5 2 9 6 9	0.01 0.004 0.01 0.33 0.13 0.14	.165-08 .110-07 .765-04 .410-08 .110-04	xxx xxx xxx xxx xxx x	.0304 xxx .0304 xxx .0304 xxx .0304 xxx .0304 xxx .0304 xxx	
	P1	40	-						
	P2	33	0011000221 0011000222	2 2	0.01 0.001	.450-04 .950-06	x xxx	.0603 xxx .0606 xxx	
	P3	39	1100111111 0011000122 1100111111	3 2 5	0.02 0.001 0.23	.190-05 .250-06 .330-05	xx xxx xx	.0903 xxx .0513 xxx .1231 xx	
	P4	19	0011000122	2	0.002	.130-05	xx	.1052 xxx	
1,2,3,4,9,10, erw.Zensur, Prozentsatz, erh.Zensur, Prüf.art	P1 - P4	131	0011001122 0011002121 1100111111 1100111122	4 3 5 4	0.006 0.02 0.39 0.12	.440-10 .600-06 .510-04 .700-05	xxx xxx - xx	.0295 xxx .0220 xxx .0341 xxx .0287 xxx	
	P1	40	-						
	P2	33	0011002121	2	0.01	.920-04	-	.0602 xxx	
	P3	39	0011001122 1100111121	2 3	0.001 0.09	.340-06 .940-04	xxx -	.0513 xxx .0749 xxx	
	P4	19	0011001122	2	0.002	.240-05	xx	.1052 xxx	

nisse für P1 weggefallen sind. Hatten wir bei der Einzelanalyse zunächst nur zwei gegensätzliche Typen gefunden (über- vs. unterdurchschnittliche erwartete, erhaltene Zensuren, Menge des bekannten Prüfungsstoffes), so erhöht sich jetzt die Zahl der auftretenden Merkmalskombination auf 5.

Gegenüber der Emotionsliste zeigt sich, daß die dort gefundenen Typen durch die Agglutination mit den Leistungsvariablen aufgespalten werden und so z. T. nicht mehr den Sprung über das alpha-adjustierte Signifikanzniveau schaffen

Durch Agglutination von Typen zeigt sich, daß die zentrale Variable für beide Datensätze die erhaltene Zensur ist. Es zeigt sich eine Beziehung zwischen einem Abschneiden mit 'sehr gut' und den Emotionen Erleichterung, Freude, Zufriedenheit, Zuversicht und Sicherheit sowie zwischen einem unterdurchschnittlichen Abschneiden mit den Emotionen Ärger und Enttäuschung.

So trivial dies manchem Leser erscheinen mag, sollte er sich aber auch über die beobachteten Häufigkeiten wundern. Wir hatten die Wahrscheinlichkeit, ein 'sehr gut' zu bekommen, auf $p(1) = .357$ geschätzt, in der Analyse über alle Datenerhebungszeitpunkte mit $N = 131$ und den Variablen der reduzierten Emotionsliste und den Leistungsdaten können wir nur 11 % der 'sehr guten' Kandidaten klassifizieren.

Wir können unsere Daten daher auch damit erklären, daß wir eine multiple Partialkorrelation vor uns haben, wobei Prüfungsstoff und erwartete Zensur

einerseits, die Zensuremotion andererseits unabhängig sind.

Diese Hypothese überprüften wir mit einem (hier nicht dargestellten) KFA-Lauf über die Items 1, 2, 3, 4, 8, 9, 10 und der erhaltenen Zensur. Die Prägnanzkoeffizienten fielen darin nicht ab.

In Tabelle 76 sind die Ergebnisse der gemeinsamen Analyse der Attribuerungsdaten und der Leistungsdaten dargestellt.

Tabelle 76: Ergebnisse der KFA mit den Leistungsdaten und den Attribuierungsskalen

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. Typ	F	ε	B	s.	Q	s.
1112212, 1212212, 2222121, 1212212, 1112212, 1212212, 2222121, 1212212	01 - 06	131	1112212	10	2,74	.130±03	x	.0556	xxx
			1212212	10	2,60	.110±03	x	.0176	xxx
			2222121	6	0,45	.680±05	xxx	.1401	xxx
			1212212	6	0,76	.110±03	x	.1174	xxx
		40 33 30 26	- - - -	- - - -	- - - -	- - - -	- - - -	- - - -	- - - -
1112212, 1212212, 2222121, 1212212, 1112212, 1212212, 2222121, 1212212	01 - 06	131	1112212	9	1,69	.140±03	x	.1064	xxx
			1212212	3	0,12	.270±03	-	.0100	xxx
			2222121	6	0,55	.140±03	xx	.0700	xxx
		40 33 30 26	- - - -	- - - -	- - - -	- - - -	- - - -	- - - -	- - - -
			-	-	-	-	-	-	-

Allgemein zeigt sich gegenüber den Einzelanalysen ein Prägnanzverlust, obwohl wir Attribuerungskollektive finden, die bei der Einzelanalyse noch nicht als Überzufällig konstatiert werden konnten.

Bei der Analyse über alle Datenerhebungszeitpunkte finden wir zwei Kollektive, die unterdurchschnitt-

lich abschnitten, dies auch erwarteten und einen unterdurchschnittlichen Prozentsatz des Prüfungsstoffes skalierten. Das eine Kollektiv attribuiert im positiven Sinne dabei ausschließlich external-variabel, das andere sowohl external-variabel als auch external-konstant. Das Kollektiv mit überdurchschnittlichen Leistungsdaten attribuiert allgemein positiv. Es zeigt sich, daß 90 % der ausschließlich external-variabel attribuirenden Unterdurchschnittlichen VDn sind, während der Typ der positiv attribuirenden Überdurchschnittlichen aufgespalten wird.

In Tabelle 77 sind die Ergebnisse der gemeinsamen Analyse über alle Datenerhebungszeitpunkte der skalierten Zensurkonsequenzen und der Attribue-rungsdaten dargestellt. In der unteren Hälfte fin-den sich die Ergebnisse eines Wiederholungslaufes unter Hinzufügung der Variablen Prüfungsart, wobei jedoch aus programmtechnischen Gründen die Werte für die Zensurkonsequenzen 9 und 10 ('Zuversicht' und 'Sicherheit') unberücksichtigt blieben.

Tabelle 77 : Ergebnisse der KFA mit den Items der emotionalen Reaktion auf die Zensur und den Kausalattributionsskalen

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. Typ	F	E	B	s.	D	s.
1/2/3/4/9/10, IK,EK,IV,EV	P1 - P4	140	0011001111	4	0.007	.770-10	xxx	.0295	xxx
			0011001211	2	0.007	.260-04	x	.0142	xxx
			0011002111	2	0.007	.240-04	x	.0142	xxx
			0011002121	2	0.005	.120-04	x	.0143	xxx
			1100111212	7	0.54	.150-05	xx	.0463	xxx
			1100112222	6	0.40	.390-05	xx	.0401	xxx
1/2/3/4, IK,EK,IV,EV,Prüf.art	P1 - P4	140	0011--11112	3	0.005	.250-07	xxx	.0214	xxx
			0011--21112	2	0.01	.160-04	xx	.0142	xxx
			0011--21211	2	0.01	.350-04	x	.0142	xxx
			1100--12121	13	2.73	.420-05	xx	.0748	xxx
			1100--22222	8	0.97	.700-05	xx	.0506	xxx

Insgesamt finden wir sechs Typen, die alle nur schwach prägnant ausfallen. Zwei Kollektive reagieren auf ihr Abschneiden mit Erleichterung, Freude, Zuversicht und Sicherheit, das eine attribuiert auf allen Dimensionen positiv, das eine auf den externalen positiv, auf den internalen negativ. Dazu finden sich die beiden Gegentypen i. e. zwei Kollektive, die mit Ärger und Enttäuschung reagieren, das eine attribuiert auf allen Dimensionen negativ, das andere auf den internalen positiv, auf den externalen negativ. Ferner finden sich noch zwei Kollektive, die mit Ärger und Enttäuschung reagieren, wobei das eine external-konstant positiv, auf den anderen Dimensionen negativ, das andere internal-konstant positiv, auf den anderen Dimensionen negativ attribuiert.

Bei Hinzufügung der Variablen Prüfungsart zeigt sich, daß sich die absolute Fallzahl für die positiv Empfindenden erhöht, daß einzelne Versuchspersonen, die nur mit Freude und Erleichterung reagierten, zusätzlich klassifiziert werden.

Wir gehen deshalb sofort zu Tabelle 78 über, in der die Ergebnisse der gemeinsamen Analyse der Leistungsdaten, der Attribuiierungsdaten und der Zensurkonsequenzen dargestellt sind.

Tabelle 78: Ergebnisse der KFA mit den Moderatorvariablen der erwarteten und erhaltenen Zensur, den Items der Zensurkonsequenzen und den Attribuierungsskalen

Variablen in der KFA	Datenerh. zeitpkt.	N	signif. Typ	F	E	B	s.	Q	s.
erwartete Zensur, erhaltene Zensur, 1/2/3/4, IK, EK, IV, EV	P1 - P4	131	1110o2222 12oo112111 22oo111111 2211oo1212	8 2 2 9	0.41 0.005 0.006 1.43	.110-o7 .120-o4 .180-o4 .160-o4	xxx x x x	.0581 xxx .0152 xxx .0152 xxx .0584 xxx	
erhaltene Zensur, 1/2/3/4, IK, EK, IV, EV	P1 - P4	140	111oo2222 2oo111111 2oo111211 2oo112111 2oo112121	13 4 2 2 2	1.05 0.01 0.01 0.01 0.008	.740-1o .540-o9 .680-o4 .650-o4 .320-o4	xxx xxx x x x	.0859 xxx .0285 xxx .0142 xxx .0142 xxx .0142 xxx	
erhaltene Zensur, 1/2/3/4, IK, EK, IV, EV, Prüf.art	P1 - P4	140	111oo2222 2oo1121211 2oo1121112 2oo1111112	8 2 2 3	0.34 0.005 0.004 0.003	.270-o8 .150-o4 .670-o5 .680-o8	xxx x xx xxx	.0549 xxx .0142 xxx .0143 xxx .0214 xxx	

Im obersten Drittel (in dem erwartete und erhaltene Zensur verrechnet wurden,) finden wir für die gemeinsame Analyse über alle Datenerhebungszeitpunkte vier Typen. Wir finden ein mäßig prägnantes Kollektiv, das eine überdurchschnittliche Zensur erwartet und erhält, mit Erleichterung und Freude reagiert und das Abschneiden auf alle vier Dimensionen zurückführt. Ferner finden wir ein mäßig prägnantes Kollektiv, das eine unterdurchschnittliche Zensur erwartet und erhält, dennoch Freude und Erleichterung empfindet und auf den externalen Dimensionen positiv, auf den internalen negativ attribuiert. Weiterhin finden sich zwei schwach prägnante Kollektive, die Ärger und Enttäuschung empfinden, wobei das eine eine überdurchschnittliche Zensur erwartet, aber nicht erhält, und dies auf internal-konstante Ursachen zurückführt und als Begründung die anderen Mög-

lichkeiten ablehnt (oder aber: die Behauptungen, die zur Berechnung der internal-konstanten Scores verwendet werden, trotzdem aufrechterhält und das Abschneiden auf das Fehlen internal-variabler und/oder external-konstanter und/oder external-variabler Umstände zurückführt), das andere eine unterdurchschnittliche Zensur erwartet und erhält und dies negativ auf alle Attribuierungsdimensionen zurückführt.

Im mittleren Drittel sind die Ergebnisse dargestellt, die sich bei einem Wiederholungslauf ergaben, wobei jedoch die erwartete Zensur unberücksichtigt blieb. Die beiden Kollektive, die auf allen Dimensionen gleichsinnig attribuieren, gewinnen dabei an Prägnanz. Das Kollektiv der external-positiv Attribuierenden verfehlt nunmehr die Signifikanzgrenze. Das Kollektiv derjenigen, die internal-konstant attribuieren, bleibt erhalten, verliert jedoch an Prägnanz. Ferner finden sich zwei neue schwach-prägnante Kollektive, die beide unterdurchschnittliche Zensuren erhalten, mit Ärger und Enttäuschung reagieren, das eine jedoch external-konstant positiv, ansonsten negativ, das andere auf den internalen Dimensionen positiv, auf den externalen negativ attribuiert.

Bei Hinzufügen der Variablen Prüfungsart finden sich nur noch vier Typen. Das Kollektiv der überdurchschnittlich guten, auf allen Dimensionen positiv Attribuierenden, Erleichterten und Erfreuten zerfällt, was zu einem Sinken des Prägnanzkoeffizienten führt, wobei jedoch das Subkollektiv der V_{Dn} die Signifikanzgrenze verfehlt. Die Prägnanz

des Kollektivs der 'Schlechten', auf allen Dimensionen negativ attribuierender Verärgerter und Enttäuschter wird bei gleicher Prägnanz der Gruppe der HDn zugeordnet. Das gleiche geschieht mit dem Kollektiv der trotz unterdurchschnittlichen Abschneidens internal-konstant Attribuierenden, während die auf beiden internalen Dimensionen positiv Attribuierenden der Gruppe der VDn zugeordnet werden. Das Kollektiv der external-konstant positiv Attribuierenden 'verschwindet'.

Zusammenfassend erhalten wir folgende Ergebnisse: Eine überdurchschnittliche erhaltene Zensur löst bei den Betroffenen die Gefühle Erleichterung, Freude, Zufriedenheit, Zuversicht und Sicherheit aus, eine unterdurchschnittliche erhaltene Zensur steht im Zusammenhang mit den Gefühlen Ärger und Enttäuschung. Überdurchschnittliche Leistungsdaten führen zur Attribuierung simultan auf Fähigkeit, Glück, das Prüfungssystem und Anstrengung. Dies betrifft VDn und HDn gleichermaßen. Unterdurchschnittliche Leistungsdaten und Attribuierungsskalen ergeben ein Kollektiv von VDn, das auf Glück (!), Faulheit, Unfähigkeit und schlechtes Prüfungssystem attribuiert, und ein Kollektiv von VDn und HDn, das auf Glück, gutes Prüfungssystem, Faulheit und Unfähigkeit attribuiert. Wenn HDn mit Erleichterung und Freude auf die Zensur reagieren, dann attribuieren sie simultan auf Glück, Anstrengung, gutes Prüfungssystem und Fähigkeit; wenn VDn mit Erleichterung und Freude reagieren, attribuieren sie auf Glück, gutes Prüfungssystem, Unfähigkeit und Faulheit. Wenn HDn mit Ärger und Enttäuschung reagieren, dann attribuieren sie das

Ergebnis entweder auf ihre Fähigkeit (!) oder ihre Unfähigkeit, in jedem Fall aber simultan auf schlechtes Prüfungssystem, Faulheit und Pech; wenn VDN mit Ärger und Enttäuschung reagieren, attribuieren sie die Zensur auf Fähigkeit, Anstrengung, schlechtes Prüfungssystem und Pech.

Es zeigt sich, daß überdurchschnittliche erwartete und erhaltene Zensur zu Erleichterung und Freude und zu simultan positiver Attribuierung auf allen Dimensionen, unterdurchschnittlich erwartete und erhaltene Zensur zu Ärger und Enttäuschung und simultan negativer Attribuierung auf allen Dimensionen führt. Unterdurchschnittlich erhaltene aber überdurchschnittlich erwartete Zensur führt zu Ärger und Enttäuschung und zur Attribuierung auf Fähigkeit (!), Pech, Faulheit und schlechtes Prüfungssystem. Ferner finden wir noch ein Kollektiv, das unterdurchschnittliche Zensuren erwartet und erhält, mit Erleichterung und Freude reagiert und dabei auf gutes Prüfungssystem, Glück, Unfähigkeit und Faulheit attribuiert.

5 DISKUSSION

5.1 Diskussion der methodischen Grundlagen

Bevor wir uns der Interpretation der Ergebnisse zuwenden, wollen wir uns kurz den methodischen Voraussetzungen für ihre Erhebung widmen. Unter methodischen Gesichtspunkten unterscheidet sich unsere Arbeit in mancher Hinsicht von anderen.

Wir wollen vier Gesichtspunkte gesondert diskutieren: die Selektion der Stichprobe, die Frage der internen Validität, die mit der retrospektiven Datenerhebung einhergehende Problematik und die Verwendung der KFA als Hauptanalysemethode.

Die Versuchspersonen, denen wir unsere Daten verdanken, stellen eine hoch selegierte Stichprobe dar. Die Laufbahn der von uns untersuchten Studenten enthielt so viele Selektionsmechanismen - Abitur, Numerus Clausus, Studienplatzlotto usw. -, daß wir mit Sicherheit keine klassische Zufallsstichprobe vor uns haben (vgl. Abbildung 6, S. 164).

Die Selektionsmaßnahmen durch den Autor beschränken sich darauf, zu einem bestimmten Termin Daten erheben zu wollen; insofern hätte jeder, der die Selektionsmechanismen bis zum Kandidaten in Prüfungen der Psychologie überstanden hat, die Chance gehabt, an der Untersuchung teilzunehmen. Keine Versuchsperson nahm an der Erhebung teil, weil eine andere Person sich (an der Untersuchung) beteiligte oder nicht beteiligte. Keine Versuchsperson nahm an der Erhebung nur deshalb teil, weil die Erhebung stattfand oder weil

wir die Daten benötigten. Systematisch unberücksichtigt blieben zwei Prüfungskandidaten: der eine war - nicht nur nach Ansicht des Autors - der deutschen Sprache kaum mächtig und der andere war - nicht nur nach Ansicht des Autors - zu stark medikamentös beeinflusst. Die Festlegung des Termins für die Datenerhebung wurde nicht davon beeinflusst, daß ein bestimmter Kandidat an der Prüfung teilnahm oder nicht teilnahm. Die Beziehung zwischen dem Autor und den konkreten Versuchspersonen ergab sich daher - aus der Sicht des Autors - rein zufällig.

Einige der Prüfungskandidaten, die die Chance gehabt hätten, an der Untersuchung teilzunehmen, verweigerten die aktive Teilnahme. Diese Gruppe der Verweigerer unterscheidet sich von den Teilnehmern dadurch, daß sie zu Beginn der Prüfung aufgeregter ist.

Wir fanden nun, daß die fremdskalierte Variable 'Aufregung zu Beginn' innerhalb der Klasse der Teilnehmer mit den Leistungsdaten negativ assoziiert ist und innerhalb der Subklasse der VDN positiv mit den Emotionstests. Dies könnte bedeuten, daß diesem einen signifikanten Unterschied zwischen Verweigerern und Teilnehmern evtl. eine größere und weitergehende Bedeutung zukommt, daß sich die Verweigerer von den Teilnehmern noch in anderer Hinsicht unterscheiden. Wir hätten damit einen weiteren Selektionsmechanismus bei der Stichprobenbildung gefunden, der den Aspekt der Teilnahmewilligkeit betrifft.

Dieser Aspekt betrifft aber nur die Selektion der Stichprobe. Der Prozeß der Datenerhebung oder die Prüfungen selbst wurden nicht berührt, da weder der Prüfer während der Prüfung noch der Beisitzer beim Ausfüllen des Beisitzerblatts wissen konnten, ob sie einen 'Teilnehmer' oder einen 'Verweigerer' vor sich hatten.

Wir wissen demnach, daß wir es mit einer hochselektierten Stichprobe zu tun hatten, die sich zudem von den Verweigerern unterscheidet. Dieser Verstoß gegen das Prinzip der Zufallsstichprobe betrifft jedoch nicht nur unsere Arbeit, sondern gilt für alle empirischen Forschungen, in denen Studenten als freiwillige Versuchspersonen verwendet werden. McDAVID & HARARI (1968, S. 403) bezeichnen denn auch den Studenten als die "weiße Ratte" der Humanpsychologie. In genau dem Ausmaß, in dem andere Forscher heuristisch von der Generalisierbarkeit ihrer Ergebnisse ausgehen, wollen wir dies auch tun.

Alle formalen Variablen, mit denen wir unsere Stichprobe beschreiben könnten, erbringen - gerade für die Unterschiede zwischen VDn und HDn - größtenteils triviale Ergebnisse: HDn sind - verglichen mit VDn - im Durchschnitt älter, haben mehr Semester - auch effektiv - studiert, haben mehr Prüfungen an einer Universität bis zum Zeitpunkt der Datenerhebung abgelegt, und zwar auch solche, die in Arbeitsaufwand und Schwere vergleichbar waren; zudem glauben HDn signifikant häufiger, einem Prüfer durch schriftliche Arbeiten aufgefallen zu sein,

(was nicht verwunderlich ist, da die HDn eine Diplomarbeit geschrieben haben, VDn hingegen durch die geltende Prüfungsordnung nicht zu schriftlichen Leistungen gezwungen sind, die notwendig zu einem Kontakt mit einem Professor führen).

Relevanz kommt diesen Ergebnissen allenfalls insofern zu, als wir nunmehr eine systematische Verzerrung unserer Daten durch systematische Falschantworten ausschließen können. Für die Interpretation der Kerndaten wichtig sind die fehlenden Mittelwertsunterschiede in den Angaben zu der Zahl der Prüfungen im Abitur sowie zur Zahl der verplemperten und beurlaubten Semester sowie zu den Angaben der Kandidaten, ob sie einem Prüfer bereits durch Seminardiskussionen oder politische Ansichten aufgefallen sind. Wir sehen darin die Berechtigung, Unterschiede zwischen VDn und HDn auf die Prüfungsordnung und damit auf das Immunisierungsparadigma zurückzuführen: HDn unterscheiden sich von VDn dadurch, daß sie mit Sicherheit mehr Erfahrungen mit mündlichen Prüfungen an Universitäten gesammelt haben und deshalb die Wahrscheinlichkeit von Kontrollkognitionen bei ihnen größer ist.

Aussagen über Besonderheiten unserer Stichprobe lassen sich auch aus Ergebnissen der peripheren Datensätze ableiten. Zugleich sollten die peripheren Datensätze als parallele Verhaltensstichproben zur Kontrolle des pauschalen Fehlervarianzrisikos dienen.

Vierzehn Tage vor der ersten mündlichen Prüfung

bearbeiteten die Kandidaten die Freiburger Beschwerdeliste (FBL). Als Ergebnis erhielten wir, daß Kandidaten in mündlichen Prüfungen höhere Scores in den Subskalen 1 (Allgemeinbefinden), 6 (Anspannung) und 8 (Schmerzen) erreichen; sie unterscheiden sich demnach von der Allgemeinbevölkerung durch eine Symptomatik, wie sie SCHULTE & TÖLLE (1973) als 'psychovegetatives Erschöpfungssyndrom' beschreiben. Von der zweiten Eichstichprobe (Studenten) unterscheiden sich Prüfungskandidaten durch höhere Werte in den Herz-Kreislauf-Magen-Darm-Items: für diese Items können wir den Begriff der 'vegetativen Dysregulation' von SCHULTE & TÖLLE (1973) übernehmen. Insgesamt gesehen finden wir also bei Prüfungskandidaten schon vor Beginn der mündlichen Prüfungen Zeichen vegetativer Labilität und unspezifischer Streß-Reaktionen. Vermutlich sind diese Ergebnisse auf den Zeitpunkt der Datenerhebung zurückzuführen. Die Distanz zu den Prüfungen ist so gering (vierzehn Tage), daß die Kandidaten ihre gesamte Lebensweise bereits auf die heranrückenden Termine abgestellt haben: sie arbeiten überdurchschnittlich lange und haben vermutlich Kompensationsmöglichkeiten wie körperliche Bewegungen oder soziale Aktivitäten eingeschränkt, was u. U. mit einer gewissen sozialen Isolation einhergeht. Welche Mechanismen im einzelnen auch immer zur beobachteten diffusen Symptomatik führen mögen, so können wir doch selbst unter den studentenfreundlichen Prüfungsbedingungen des Bochumer Psychologischen Instituts bereits im Vorfeld der Prüfungen Zeichen von Prüfungsbelastungen feststellen. Für die Interpretation der Kern-daten folgt daraus, daß die Prüfungsbelastung im

engeren Sinne - also die Belastungen während der Prüfung - nicht unbedingt ausschließlich an die gerade wirksamen Reize gekoppelt werden muß, sondern evtl. nur eine Intensivierung einer allgemeinen Labilität darstellt.

In der Analyse des Arbeitsverhaltens hatten wir sechs Faktoren gefunden:

1. Unsicherheit bzw. Sicherheit in der Arbeit,
2. Stetigkeit und Konzentriertheit bzw. Unstetigkeit und Ablenkbarkeit in der Arbeit,
3. allgemein adäquates bzw. allgemein inadäquates Arbeitsverhalten,
4. soziales Arbeitsverhalten,
5. konkretes adäquates bzw. inadäquates Arbeitsverhalten,
6. Impulsivität und Zerfahrenheit.

Die Dimensionen der Unstetigkeit (bzw. Stetigkeit) und der Unsicherheit (bzw. Sicherheit), wie sie MOELLER & SCHEER (1970) gefunden haben, konnten wir überzeugend replizieren. Abweichungen und Unterschiede in unserer Faktorenlösung können zum einen auf die Größe unserer Stichprobe im Verhältnis zur Itemzahl (41 Versuchspersonen und 50 Items) zurückgeführt werden, aber auch inhaltlich auf folgende drei Varianzquellen:

1. MOELLER & SCHEER (1970) berichten über fachspezifische Unterschiede, wonach Psychologiestudenten wesentlich unregelmäßiger arbeiten als Studenten anderer Fächer. Inwieweit das auch für Prüfungskandidaten zutrifft, bleibt offen. Es kann sein, daß unsere Lösung ein Resultat spezifischer Arbeitsbedingungen einer selektierten Fachgruppe ist.

2. Unsere Datenerhebung fand ca. vierzehn Tage vor Beginn der jeweils ersten Prüfung statt. Es ist nicht ausgeschlossen, daß die Kandidaten sich zu diesem Zeitpunkt bereits auf spezifische Prüfungsfächer vorbereiteten und prüfungsspezifisches Arbeitsverhalten realisierten. SUAREZ (1971) beispielsweise demonstrierte an Mittelschülern, daß in jeweils anderen Fächern jeweils andere Arbeitstechniken zum Erfolg führten.
3. Ferner ist anzunehmen, daß verschiedene Prüfungsformen sich unterschiedlich auf das Arbeitsverhalten auswirken (vgl. GAUDRY & BRADSHAW, 1970; BALCH, 1971; KVALE, 1972, S. 69f., BOYLE, 1969). Zu berücksichtigen ist daher, daß sich die Kandidaten unserer Stichprobe auf mündliche Prüfungen vorbereiteten, was einen Vergleich mit 'Normal'-Studenten erschwert.

Interessant an unserer Lösung ist, daß wir erstmals auf einen Faktor sozialen Arbeitsverhaltens gestoßen sind. Die Items, die auf diesem Faktor laden, thematisieren den Kontrast Gruppen- vs. Einzelarbeit. Diesem Aspekt ist in der bisherigen Forschung wenig Aufmerksamkeit gewidmet worden, und wir sind die ersten, die ihn faktorenanalytisch sichern konnten. Insgesamt gesehen gelang es also, die beiden wichtigsten Dimensionen zu replizieren. Die anderen von uns gefundenen Dimensionen können als methodisches Artefakt aufgefaßt werden, aber auch als Resultat des Studiengangs der untersuchten Stichprobe, der Vorbereitung einer größeren Gruppe auf ein spezifisches Prüfungsfach oder als Folge der Vorbereitung auf mündliche Prüfungen.

Mit insgesamt sieben Analogskalen untersuchten wir die Emotionsverläufe. Wir ließen die Angst, das Lampenfieber, die Depressivität, die Hilflosigkeit, die Wut, den Ärger und die Freude retrospektiv 'vor'

und 'während' der Prüfung und on-line nach der Prüfung skalieren. Für die methodischen Grundlagen der Kerndaten war der Nachweis wichtig, daß retrospektive Messungen von Emotionen (auch wenn sie in der Prüfungsforschung der Standard sind) dann valide Daten liefern, wenn man sie unmittelbar nach dem emotionsauslösenden Ereignis vornimmt; wichtig war ferner die Replikation des Angstverlaufs.

Die Ergebnisse haben wir in Kapitel 4.2.3 dargestellt. Eine Reduktion der Intensität können wir für die Skalen 'Angst' und 'Lampenfieber' nachweisen, wobei sich zwischen den semantisch ähnlichen Skalen ein Unterschied zeigt: 'Angst' geht stärker vom Zeitpunkt 'vor' der Prüfung zum Zeitpunkt 'während' der Prüfung zurück, 'Lampenfieber' stärker von der 'Während-Messung' zur On-line-Messung 'nach der Prüfung' zurück.

Wir betrachten dies als Replikation bisheriger Forschungsergebnisse.

Die Ergebnisse zum Angstverlauf sind um so überzeugender, als die Emotion der 'Freude' einen genau entgegengesetzten Verlauf zeigt; die 'Freude' ist am niedrigsten 'vor' der Prüfung und steigt dann kontinuierlich an. 'Freude' und 'Angst' betrachten wir im Zusammenhang mit Prüfungen als inkompatible Emotionen. Wir unterstellen dabei, daß eine Emotion wie 'Freude an der Angst', wie sie bei Horrorfilmbesuchern zu beobachten sein mag, in Prüfungen nicht zu beobachten ist. 'Freude' steigt nach unseren Ergebnissen von einem Tiefpunkt 'vor'

der Prüfung her kontinuierlich an, 'Angst' und 'Lampenfieber' erreichen ihr Maximum unmittelbar 'vor' der Prüfung und fallen dann ab.

Für die 'Depressions'-Skala erhalten wir Ergebnisse, die dort, wo sie signifikant ausfallen, ebenfalls einen abfallenden Verlauf zeigen.

Insgesamt gesehen erhielten wir keine Ergebnisse, die auf eine Interaktion hindeuten. Dies bedeutet, daß der Unterschied zwischen erfahrenen und unerfahrenen Fallschirmspringern (vgl. FENZ & EPSTEIN 1968) nicht als Unterschied zwischen erfahrenen und - wahrscheinlich - unerfahrenen Prüfungskandidaten repliziert werden konnte. Bei der Diskussion der Kerndaten werden wir freilich sehen, daß der Unterschied zwischen VDN und HDN dem Unterschied zwischen erfahrenen und unerfahrenen Fallschirmspringern auch nicht entspricht.

Bislang sind wir auf die besonderen Wirkungen einer retrospektiven Messung nur insofern eingegangen, als wir die Replikationen der Ergebnisse anderer Forscher durch eine Mischung von retrospektiven und On-line-Messungen aufgezeigt haben.

Diese Replikationen sprechen zumindest gegen die Annahme, daß retrospektive Messungen unmittelbar nach dem Ereignis, auf das sich die Messungen beziehen, grobe Verzerrungen mit sich bringen.

In unserem Design haben wir den größten Teil unserer Daten nach Bekanntgabe der Zensur erhoben. Von

daher ergibt sich die Möglichkeit, daß die 'Kenntnis der Zensur' und damit einhergehende Reaktionen wie Freude über ein unerwartet gutes Abschneiden oder Ärger über eine ungerechte Zensur die retrospektiven Messungen verändern. Wir haben versucht, dies durch On-line-Messung der Emotionstests in der Baseline und durch retrospektive Messungen nach den Prüfungen in den Griff zu bekommen (in beiden Fällen errechneten wir die gleiche multivariate Struktur).

Unser Ergebnis, daß bei den Analogskalen innerhalb einer 'vor'-'während'-'nach'-Messung sich jeweils Kollektive über- und unterdurchschnittlicher Merkmalsträger zusammenfinden, kann bei den Skalen 'Depressivität', 'Wut' und 'Ärger' auf hohe Merkmalsstabilität zurückgeführt werden; dies kann natürlich auch mit dem Erinnerungseinfluß einer immer gleich guten oder schlechten Zensur begründet werden, nur ergibt sich dabei die Schwierigkeit, daß ein immer gleicher Erinnerungseinfluß sich nicht in einer hohen Merkmalsstabilität der Skalen 'Angst', 'Lampenfieber' und 'Freude' niederschlägt. Somit müssen wir zwei Erklärungsmöglichkeiten auf ihren logischen Gehalt untersuchen und gegeneinander abwägen: zum einen die Möglichkeit, daß Erinnerungseinflüsse die einen Skalen betreffen, die anderen nicht, zum anderen die Möglichkeit, daß die Erinnerungseinflüsse so geringfügig sind, daß sie vernachlässigt werden können. Die logisch einfachere Möglichkeit ist u. E. die letzterwähnte; sie ist daher vorzuziehen.

Insgesamt gesehen bietet sich in unserem reichhaltigen Datenmaterial kein Hinweis darauf, daß der Akt der Bekanntgabe der Zensur die retrospektive Messung in einem Ausmaß 'verfälscht', der berücksichtigt werden müßte.

Zum Teil mag dies auch daran liegen, daß wir non-parametrisch mit der KFA gearbeitet haben, wobei sich minimale Verzerrungen wegen der Dichotomisierungen nicht 'durchsetzen' können.

Damit kommen wir zu einem Punkt der Diskussion, der die Kerndaten von einem methodisch-statistischen Standpunkt her tangiert. Unseres Wissens legen wir die erste Arbeit vor, in der die KFA als Methode der Wahl in einem solchen Umfang eingesetzt wurde.

Ein Nachteil der KFA ist ihre Unhandlichkeit. Die betrifft nicht nur die - computertechnisch bedingte - Beschränkung auf maximal zehn berechnete Variablen.

Wenn man auf den Chi-Quadrat-Ansatz der KFA verzichtet und über den Ansatz mit simultanen Binomialtests rechnet, so umgeht man zwar das Problem der ausreichenden Zellbesetzung (i. e. ausreichende Zahl von Versuchspersonen), man hat sich aber gleichzeitig ein neues Problem eingehandelt: den Verzicht auf heuristische Suchverfahren und Exklusionsprinzipien.

Dies heißt zunächst, daß andere statistische Verfahren in den Fällen eingesetzt werden müssen, in

denen es um die Auswahl typologisch relevanter Variablen geht. Daraus ergeben sich zwei Folgeprobleme.

Das erste Folgeproblem betrifft die zugrunde gelegte Modellvorstellung. Der routinemäßige Einsatz der Faktorenanalyse als eines Verfahrens, das Fehlervarianz eliminiert und die Information auf wenige 'repräsentative' Items reduziert, ist nicht immer angemessen. Unsere Analysen des Testverhaltensdifferentials und vor allem der Strategiensammlung demonstrieren überzeugend, daß 'Routine'-Faktorenanalysen in die Irre führen können. Mancher Autor wäre infolge der Faktorenlösungen vielleicht zu dem Schluß gekommen, die Datensätze enthielten in erster Linie Zufallszahlen und seien somit unbrauchbar. Tatsächlich zeigt sich aber in den weiteren Analysen, daß lediglich das Modell der Faktorenanalyse bei diesem Datensatz ungeeignet ist, weil - bei der Strategiensammlung besonders deutlich - zwei Kollektive mit reliablen Strategien dem Rest der Versuchspersonen gegenüberstehen, die stark fluktuierend skalieren. Mit der KFA können wir diese Ergebnisse berechnen. Dadurch jedoch, daß wir gezwungen sind, vor die KFA andere Verfahren zu setzen, erhöht sich das Beta-Risiko, wenn wir nicht klare Modellvorstellungen von den zu erwartenden Ergebnissen haben. Unsere Modellvorstellung von Strategien beinhaltet die Möglichkeit adimensionaler Strukturen und den Kontrast 'Kontrollstrategie' vs. 'Gelernte Hilflosigkeit', wobei sich der Kontrast möglicherweise nicht als niedrige und hohe oder positive und negative Ausprägung zeigt,

sondern evtl. als Kontrastassoziation zwischen Items vs. Zufallszahlen. Unsere Modellvorstellung war also relativ unpräzise, weil fast alle Alternativen zur Modellvorstellung der Faktorenanalyse aufgenommen wurden. Die Folge davon war, daß wir aufwendige und umfangreiche Voranalysen vornahmen, bis wir dann schließlich einen kleinen Satz von Variablen in der KFA verrechnen konnten.

Dies führt zum zweiten Folgeproblem. Wir haben denselben Datensatz mehrfach verrechnet; dies tangiert das nominelle Alpha-Risiko. In der KFA ist deshalb die Maßnahme der Alpha-Adjustierung vorgesehen. In unserem Fall jedoch haben wir einzelne Datensätze so häufig verrechnet, daß eine präzise Alpha-Adjustierung nicht mehr möglich war. Wir haben deshalb selbst in den Fällen zur Maßnahme der totalen Alpha-Adjustierung gegriffen, in denen dies wegen präziser Hypothesen nicht nötig gewesen wäre. Dies bedeutet, daß wir in einzelnen Datensätzen das Beta-Risiko unnötig erhöhten, in anderen Datensätzen möglicherweise das Alpha-Risiko nicht ausreichend berücksichtigten.

Unsere technisch bedingte Vorgehensweise, vor der KFA andere Verfahren einzusetzen, führt - wie wir aufzuzeigen versuchten - zu methodisch-statistischen Folgeproblemen, die bislang in Lehrbüchern kaum oder nicht berücksichtigt wurden. Unseren Ergebnissen haftet deshalb eine gewisse Heuristik an.

Dies betrifft auch die Agglutination der einzelnen Datensätze, die ebenfalls heuristisch zu sehen sind.

Hier haben wir die logische Überlegung, daß eine definierte Menge von Merkmalsträgern mal mit dem einen Datensatz, mal mit einem anderen Datensatz erfaßt wird, in Plasmoden simuliert und Agglutinations-KFAn gerechnet. Ebenso simulierten wir den Fall der Agglutination zweier unabhängiger Datensätze, wobei wir mit jedem einzelnen Datensatz jeweils hochsignifikante Typen errechnen: Bei der Agglutination erhielten wir durchaus 'signifikante' Typen, deren Prägnanzverlust gegenüber den Einzelanalysen jedoch drastisch ausfiel. Unter Berücksichtigung dieser Simulationsergebnisse entwickelten wir das in Kapitel 3.5.3 dargestellte heuristische Beurteilungsverfahren.

Insgesamt gesehen beruhen unsere Ergebnisse also auf einer Reihe von heuristischen Annahmen, die wir wegen des Fehlens präziser mathematisch-statistischer Modelle treffen mußten. Insofern kann die Interpretation unserer Ergebnisse nur 'vorläufig' sein. Dies ist ein Nachteil der KFA, den wir uns durch ihren Einsatz einhandelten. Dem steht der Vorteil gegenüber, daß wir Ergebnisse erzielten, die mit anderen Verfahren nicht oder nur schwer möglich waren.

Beispiele dafür bietet die Analyse der Strategiensammlung. Auch die in Kapitel 4.3.1 dargestellte Assoziation zwischen erwarteter und erhaltener Zensur und dem Prozentsatz des geprüften Stoffes, mit dem sich der Kandidat während des Studiums beschäftigt hat, bietet ein Beispiel dafür, daß mit der KFA logischere Ergebnisse erzielt werden können: In der parametrischen Analyse fanden wir eine sig-

nifikante positive Korrelation zwischen erwarteter und erhaltener Zensur und keine bedeutsame Korrelation zwischen diesen beiden Variablen und dem Prozentsatz des während des Studiums bekannt gewordenen Stoffes, was inhaltlich als Unabhängigkeit von Prüfungsstoff und Studieninhalten zu interpretieren wäre; in der KFA finden wir einen Typ, der eine überdurchschnittliche Zensur erwartet und erhalten wird und von daher einen Zusammenhang mit einem überdurchschnittlichen Prozentsatz des Prüfungsstoffes, der Inhalt des Studiums war, herstellt, und einen Gegentyp, der eine unterdurchschnittliche Zensur erwartet und erhalten wird und dies in Beziehung zu einem unterdurchschnittlichen Prozentsatz des geprüften Stoffes setzt, der ihm bereits im Studium begegnet ist. Inhaltlich heißt dies, daß diejenigen Prüfungskandidaten, die ihre Zensur korrekt als über- oder unterdurchschnittlich antizipierten, sich auf das Ausmaß beziehen, in dem vor Beginn der Prüfungsvorbereitung bearbeitetes (oder nicht bearbeitetes) Stoffgebiet Gegenstand der mündlichen Prüfung war. Dieses Ergebnis stimmt überein mit unserem anderen Ergebnis, daß über- und unterdurchschnittliches Abschneiden in einem Zusammenhang mit lernintensiven Lehrveranstaltungen in den letzten beiden Semestern steht. Das Ergebnis, das wir mit der KFA berechnen, ist also logischer!

Wie das letzterwähnte Ergebnis auch zustande kommen mag, ob hier eine Assoziation höherer Ordnung vorliegt oder ob die Typen durch Verrechnung von Extremwerten sich zusammenfinden, wir haben Grund, den Ergebnissen der KFA zu trauen.

Drei methodische Voraussetzungen sind bei der In-

terpretation der Kerndaten zu berücksichtigen: die Besonderheiten der hochselegierten Stichprobe, die Besonderheiten des statistischen Verfahrens und die Ergebnisse der Peripherdaten.

Die Ergebnisse der Peripherdaten, methodisch gesehen handelt es sich um parallele Verhaltensstichproben, sind insgesamt so überzeugend, daß wir einen pauschalen Einwand zurückweisen können, die Ergebnisse der Kerndaten könnten 'unter diesen Verhältnissen' (der Datenerhebung) und 'bei dieser Stichprobe' sowieso nicht reliabel sein, weil die Fehlervarianz zu groß sei.

5.2 Diskussion des hypothesenprüfenden Teils

Die Diskussion der Ergebnisse dieses Teils wollen wir zunächst den Einzel-Hypothesen widmen, bevor wir auf die Hypothesennetze übergehen.

Aus zwei theoretischen Modellen haben wir ein besseres Abschneiden der HDn im Vergleich zu den VDN abgeleitet. Im Modell der Kritischen Theorie wird die Institution der Universität als Organisation konzipiert, die rationalen ökonomischen Prinzipien folgt und deshalb die Selektion Ungeeigneter frühzeitig in den Zwischenprüfungen vornimmt; eine Prüfung der Hypothese 1 im Zusammenhang mit diesem Modell beschränkt sich allein auf die selektionsrelevante Variable der erhaltenen Zensur. Das von uns konzipierte Modell der Strategien sieht vor, daß der Kandidat unter Einsatz von Strategien seinen Zielzustand - eine möglichst gute Zensur - erreichen

will; eine Prüfung der Hypothese 1 nach diesem Modell kann sich nicht nur auf die erhaltene Zensur als Kriterium beschränken, sondern man muß das Erreichen des Zielzustandes differenzierter erfassen: Man kann annehmen, daß ein Kandidat, der eine überdurchschnittliche Zensur erhält und diese auch erwartet hat (wobei Prüfer und Beisitzer zu der gleichen Beurteilung gelangen wie der Kandidat), beide Zensuren zudem auf den Prüfungsstoff bezieht, tatsächlich einen Zielzustand erreicht hat, und daß ein Kandidat, der eine unterdurchschnittliche Zensur erwartet und in Übereinstimmung mit Prüfer und Beisitzer auch erhält, tatsächlich einen Zielzustand nicht erreicht hat. Der multivariate Datensatz der 'Leistungsdaten' entspricht daher eher der Operationalisierung des Strategiemodells.

Die Ergebnisse haben wir in Kapitel 4.3.1 en detail dargestellt und in Kapitel 4.3.6 überblicksartig zusammengefaßt: Wir fanden kein Indiz dafür, daß VDN häufiger mit "mangelhaft" abschneiden, als HDn. Ferner fallen sowohl die vorhergesagte multivariate Assoziation der überdurchschnittlichen Merkmalsausprägungen der Leistungsdaten mit der Merkmalsausprägung HD und - als Kontrast - die vorhergesagte Assoziation der unterdurchschnittlichen Leistungsdatenmerkmalsausprägungen und der Merkmalsausprägung VD der Variablen Prüfungsart signifikant aus. Wenn wir den multivariaten Datensatz als 'Leistung aus der Sicht des Kandidaten' auffassen, so kommen wir zur Aussage, daß Kandidaten in Abschlußprüfungen im Schnitt bessere Leistungen erreichen als Kandidaten in Zwischenprüfungen (i. e. Hypothese 1 wird angenommen).

Gleichzeitig sprechen unsere Ergebnisse mehr für das Strategiemodell und gegen das Selektionsmodell. Über die Annahme der Hypothese hinaus können wir unsere Aussagen aufgrund der KFA-Ergebnisse auch noch weiter präzisieren: Das aus der Einzelanalyse der Leistungsdaten bekannte Kollektiv der 'Überdurchschnittlichen' wird bei Hinzufügung der Variablen Prüfungsart 'aufgespalten', während das Kollektiv der 'Unterdurchschnittlichen' der Merkmalsausprägung VD zugeordnet wird. Dies bedeutet inhaltlich, daß Überdurchschnittliche Leistungsdaten kein 'Privileg' der HDn darstellen, sondern ebenso von VDn erreicht werden; umgekehrt sind unterdurchschnittliche Leistungsdaten hingegen ein signifikantes 'Privileg' der VDn. Wir werden diesen Punkt an späterer Stelle als Folge des Immunisierungsparadigmas ausführen.

Als 'emotionales Defizit' hatten wir eine Mischsymptomatik aus Angst und Depression als "Prüfungsbelastung" postuliert (Hypothese 6); zudem haben wir vorausgesagt, daß VDn stärkeren Prüfungsbelastungen unterworfen waren als HDn (Hypothese 8). Die Ergebnisse haben wir en detail im Kapitel 4.3.2 dargestellt und in den Kapiteln 4.3.2.3 und 4.3.6 überblicksartig zusammengefaßt.

Es zeigt sich, daß sich die jeweils hohen und die jeweils niedrigen Merkmalsausprägungen aller drei Tests sowohl in der On-line-Messung der Baseline als auch in den retrospektiven Messungen zu hochprägnanten Typen 'Hoch-Emotionalisierter' bzw. 'Niedrig-Emotionalisierter' zusammenfinden.

Agglutinations-KFAn mit den on-line nach der Prüfung gemessenen Analogskalen 'Erschöpfung' und 'Konzentrationsfähigkeit' sowie mit den Faktorscores der retrospektiv gemessenen Beschwerdeliste ergeben Typen mit jeweils über- und unterdurchschnittlichen Merkmalsausprägungen. Dies deutet auf konvergente Validität der Belastungsmessungen, und wir können Hypothese 6 annehmen. Prüfungsbelastungen stellen also ein Angst-, Depressions- und Hilflosigkeitsgemisch mit ausgeprägten körperlichen Symptomen dar, das nach der Prüfung mit überdurchschnittlicher Erschöpfung und verminderter Konzentrationsfähigkeit einhergeht.

Unsere Ergebnisse zeigen dabei, daß Prüfungsbelastungen von Prüfung zu Prüfung deutlich fluktuieren, wobei jedoch die Struktur der Belastungen - das 'Profil' - dadurch nicht berührt wird. Einerseits gelingt es mit einer Ausnahme nicht, in Retest-KFAn der einzelnen Belastungsvariablen Kollektive nachzuweisen, die während aller Prüfungen überdurchschnittlich ängstlich, hilflos oder depressiv sind. Andererseits erhalten wir in den KFAn der Differenzwerte signifikante und prägnante Kollektive, was auf simultane Veränderungen aller drei Emotionstests deutet; und in den Faktorenanalysen der Beschwerdeliste erhalten wir sehr ähnliche Faktormatrizen, so daß wir zwar von einer hohen Retest-Reliabilität der Faktorenlösungen, nicht jedoch von einer hohen Reliabilität der Meßwerte einzelner Versuchspersonen sprechen können. (Ähnliche Ergebnisse haben wir auch bei den Analogskalen zur Messung des Angstverlaufs gefunden: bei den Skalen Angst, Lampenfieber und

Freude gab es Ergebnisse, die auf eine gleichbleibende Ausprägung während einer einzelnen Prüfung deuten, nicht jedoch auf gleichbleibende Ausprägungen während der gesamten Prüfungszeit.)

Unsere Ergebnisse enthalten somit keinen Hinweis darauf, daß etwa eine 'Disposition zur Prüfungsbelastung' wirksam wäre. Wir haben keine Anhaltspunkte für eine Bestätigung des Stereotyps des 'Immer-überdurchschnittlich-Prüfungsbelasteten' gefunden.

Dabei zeigt sich allgemein, daß das Kollektiv der 'Niedrig-Emotionalisierten' durch Hinzufügung der Variablen Prüfungsart 'aufgespalten' wird, während die überdurchschnittlichen Merkmalsausprägungen in allen Tabellen der Merkmalsausprägung VD zugeordnet wird. Inhaltlich heißt dies, daß unterdurchschnittliche Prüfungsbelastungen kein 'Privileg' der HDn darstellen, sondern ebenso bei VDn auftreten können; umgekehrt sind überdurchschnittliche Belastungen hingegen ein signifikantes 'Privileg' der VDn.

An dieser Stelle zeigt sich besonders deutlich eine Interpretationsschwierigkeit, die mit der Anwendung der KFA und mit der Möglichkeit einhergeht, sie sowohl für Posthoc- als auch für Anthac-Analysen einzusetzen. Am deutlichsten zeigt sich diese Schwierigkeit im Kapitel 4.3.2.2 (Multivariate Analyse von H1S, STAI und BDI). Dort hatten wir folgende signifikante Typen gefunden: unterdurchschnittlich belastete VDn, überdurchschnittlich belastete VDn, unterdurchschnittlich belastete HDn. Folgt man nun dem üblichen Interpretationsalgorithmus der Varianzanalyse und interpretiert man alle signifikanten Ergebnisse en bloc, so heißt das, daß über VDn keine Aussagen gemacht werden können, da für sie sowohl über- wie unterdurchschnittlich

belastete Typen errechnet werden. Ferner muß berücksichtigt werden, daß die Konfigurationsfrequenz der (nicht-signifikanten) überdurchschnittlichen HDn in das Vertrauensintervall der (signifikanten) unterdurchschnittlichen HDn fällt. Verknüpft man also den Interpretationsalgorithmus der Varianzanalyse mit dem Algorithmus des Konfigurationsfrequenzvergleichs nach ROEDER (1976), so kommt man zur Aussage, daß von der Anthrac-Klasse nicht auf die (abhängige) Merkmalsausprägung geschlossen werden kann. Inhaltlich entspricht dies der Annahme der Nullhypothese! Legt man jedoch den strengeren Algorithmus von KRAUTH & LIENERT (1973) zugrunde, wonach vorher Hypothesen hinsichtlich Überzufällig häufiger (oder seltener) Konfigurationsfrequenzen zu formulieren sind, die allein dann überprüft werden, so kommt man in diesem Falle zur Annahme der Alternativhypothese (VDn sind überdurchschnittlich, HDn unterdurchschnittlich belastet)! In der Varianzanalyse entspräche dies dem Verzicht auf die Interpretation unvorhergesehener Signifikanzen. In der uns bekannten Methodikliteratur ist dieses Dilemma bislang nicht ausreichend gewürdigt worden. Die zukünftige Forschung erst wird daher zeigen, ob unser Entschluß, zunächst zur Hypothesenprüfung den strengeren Algorithmus anzuwenden, um die Hypothese dann (heuristisch) zu differenzieren, gerechtfertigt ist.

Wir haben somit zwei strukturell gleiche Hypothesen angenommen und strukturell gleiche Differenzierungen vorgenommen: VDn unterscheiden sich von HDn durch unterdurchschnittliche Leistungsdaten und überdurchschnittliche Belastungen, während HDn entgegengesetzte Merkmalsausprägungen (überdurchschnittliche Leistungsdaten und unterdurchschnittliche Belastungen) skalieren, dieses 'Privileg' jedoch mit einer kleinen Gruppe von VDn teilen müssen.

Als gemeinsames Ergebnis der Agglutinations-KFAn mit den Leistungsdaten und den Belastungsmessungen hatten wir hervorgehoben, daß die überdurchschnittlichen Belastungsmessungen, die unterdurchschnittlichen

Leistungsdaten und die Merkmalsausprägung VD sich zu Typen zusammenfinden, wobei wir gegenüber den jeweiligen Einzelanalysen geringfügige oder gar keine Pränanzverluste verzeichnen. Wiederum finden wir, daß die Kollektive mit den 'günstigen' Merkmalsausprägungen durch Hinzufügen der Variablen Prüfungsart 'aufgespalten' werden, was in den Agglutinations-KFAn mit den Leistungsdaten und den Emotionstests einerseits und den Faktorscores der Beschwerdenliste andererseits zur Berechnung eines signifikanten Typs mit den Merkmalsausprägungen HD, unterdurchschnittliche Belastung und überdurchschnittliche Leistungsdaten führt, für den jedoch gegenüber den Einzelanalysen Pränanzverluste verzeichnet werden müssen.

Gleichzeitig registrieren wir in der Agglutinations-KFA mit den Leistungsdaten und den Emotionstests - den zentralen Belastungsindikatoren - ein Kollektiv von VDn mit unterdurchschnittlichen Belastungsausprägungen und unterdurchschnittlichen Leistungsmerkmalen. Zugleich verfehlen manche Konfigurationsfrequenzen mit dem Merkmal HD die Signifikanzgrenze.

Insgesamt erhalten wir Ergebnisse, die inhaltlich einer starken Assoziation von unterdurchschnittlichen Leistungsmerkmalen und überdurchschnittlichen Belastungen innerhalb der Klasse der VDn entsprechen, während abweichende Ausprägungen dieser drei Datensätze weniger stark assoziiert zu sein scheinen.

Wenn wir diese Ergebnisse insgesamt auf das Immunisierungsparadigma in der Forschung zur Gelernten Hilflosigkeit übertragen, ergibt sich das im folgenden beschriebene Zwischenbild. Das Immunisierungsparadigma besteht in der Unterscheidung von zwei Gruppen, deren Vorerfahrung wir in dem einen Fall genau kennen, in dem anderen Fall nur vermuten können. In unserer Stichprobe haben HDn mit Sicherheit Erfahrungen mit Prüfungen an Universitäten gesammelt, von der Klasse der VDn wissen wir dies nicht genau: Wir vermuten, daß diese Kandidaten in ihrer Mehrheit keine Vorerfahrungen sammeln konnten, können aber nicht ausschließen, daß zu dieser Gruppe auch Studienfachwechsler oder Zweitfachstudenten mit entsprechenden Vorerfahrungen zählen. Wir untersuchten also eine Gruppe, die zu 100 % aus Erfahrenen bestand, und eine Gruppe, die überwiegend aus Unerfahrenen sich zusammensetzen sollte.

In diesem Punkt unterscheiden sich unsere beiden Gruppen von den Fallschirmspringergruppen der Autoren FENZ und EPSTEIN: sie verglichen 100%-Erfahrene mit 100%-Unerfahrenen. Unsere Ergebnisse sind deshalb nicht direkt vergleichbar!

Wir nehmen an, daß Vorerfahrungen Gelegenheiten zur Ausbildung von 'Kontrollstrategien' beinhalten. Unsere Ergebnisse sind insofern angemessen, als HDn überzufällig häufig überdurchschnittliche Leistungsdaten und unterdurchschnittliche Belastungen skalieren, VDn überzufällig häufig unterdurchschnittliche Leistungsdaten und überdurchschnittliche Belastungen berichten, gleichzeitig aber eine - vermutlich erfahrene - Gruppe von VDn das Verhaltensmuster der Erfahrenen zeigt.

Für die Strategiemessungen haben wir die Items der Strategiensammlung und die Testverhaltens-differentiale eingesetzt. Die dazugehörige Einzel-Hypothese lautete, daß es mindestens zwei multi-variate Profile gibt, wovon das eine einer Kontroll-strategie und das andere der Strategie der Gelern-ten Hilflosigkeit entspricht und wovon das eine die Umkehrung des anderen ist.

Pauschal können wir diese Hypothese annehmen, weil wir bei der Analyse der Kernstrategien den Typ der 'Assertiven' und den Gegentyp der 'Unsicheren', bei der Analyse der Trivialitems den Typ der 'Konformen' und den Gegentyp der 'Differenzierten', bei der Analyse der TVD-Versionen ebenfalls zwei Gegentypen nachweisen konnten (mäßig-angespanntes, variables, diversifizierendes Vorgehen steht einem angespannt, stetigen, unifizierenden Vorgehen gegenüber).

Der erste Schritt der Analyse der Strategien, der formale Nachweis von Typ und Gegentyp war demnach in allen drei Datensätzen erfolgreich.

Im zweiten Schritt untersuchen wir, ob die Typen inhaltlich einem in das Prüfungskommunikationssystem aktiv eingreifenden und es steuernden Verhalten und in der Umkehrung einem passiv treibenlassenden Verhalten zugeordnet werden können.

In den Analysen der Kernstrategieitems der Strategien-sammlung finden wir ein Kollektiv der 'Assertiven' (die angeben, einem Prüfer bei Meinungsverschieden-heiten auch in einer Prüfung widersprechen zu wol-len, davon ausgehen, daß ihnen in einer Prüfung

auch auf schwere Fragen eine Antwort einfallen wird, und es für besser hält, kritisch aufzutreten) und ein Kollektiv der 'Unsicheren' (die nicht widersprechen und nicht allzu kritisch auftreten wollen, zudem glauben, ihnen würde auf schwere Fragen nichts einfallen). In den Agglutinations-KFAn zeigt sich, daß die 'Assertiven' 'angespannt-stetig-unifizierend' vorgegangen sind und vorgehen wollen, die 'Unsicheren' 'mäßig angespannt-variabel-diversifizierend' vorgegangen sind und vorgehen wollen. Insgesamt erhalten wir zwei Kollektive, wovon das eine als 'selbstbewußt, aktiv und gelassen' beschrieben werden kann und seine Vorgehensweise als 'konzentriert, unbeirrbar und Zusammenhänge betonend'. Der Gegentyp stellt sich als unsicher, variabel und ablenkbar und seine Vorgehensweise als 'unbekümmert, unsystematisch und den Einzelheiten verpflichtet' dar. Wir meinen, daß Unterschiede in den Kontrollaspekten, d. h. in der Bereitschaft, ggf. in das Prüfungsgeschehen einzugreifen und von sich aus den Prüfungsverlauf zu beeinflussen, deutlich zu erkennen sind. Die Aktivitätsdimension (die theoretisch dem motivationalen Defizit in der Gelernten Hilflosigkeit entspricht) stellt sich als Unterscheidung aktiv-reaktiv dar.

Bei der Analyse der Trivialitems haben wir die Kollektive der 'Differenzierten' und der 'Konformen' gefunden. Die Items 8 ("Am Anfang bin ich sicherlich aufgeregt, aber das wird sich dann legen."), 26 ("Es kommt darauf an, Zusammenhänge zu sehen und zu beschreiben.") und 28 ("Es kommt darauf an, daß ich die Zusammenhänge und Unterschiede anschaulich darstelle.") wurden in der von uns untersuchten

Stichprobe eindeutig zustimmend beantwortet; sie erfassen somit Selbstverständlichkeiten. Das Kollektiv derjenigen, die alle drei Items simultan eindeutig mit 'richtig' beantworten, haben wir darum die 'Konformen' genannt, diejenigen, die zu nicht-eindeutigen Stellungnahmen gelangen, haben wir als 'Differenzierte' bezeichnet. Die Kollektive, die wir mit den Trivialitems errechneten, und die Kollektive aus den Kernstrategieitems überschneiden sich nur in einem Falle: die 'Unsicheren' reagieren auch 'differenziert'. 'Unsichere' verweigern also gleichzeitig Stellungnahmen auf triviale Fragen. Im übrigen scheinen die Trivialitems einen anderen Aspekt zu messen als die anderen zur Strategiemessung eingesetzten Datensätze. Inhaltlich gesehen gelingt es nur schwer, die Kollektive aus den Trivialitems auf ein aktiv kontrollierendes oder passiv treibenlassendes Vorgehen zu beziehen. Ihnen scheint mehr eine Kalkulation von diskriminativen Reizen zugrunde zu liegen.

Während Kernstrategien und Testverhaltensdifferentiale inhaltlich unschwer auf das theoretisch postulierte Verhalten bezogen werden können, ist dies bei den Trivialitems schwerer: hier müssen wir beachten, daß die mangelnde Berücksichtigung von nach allgemeinem Konsensus sicheren Ereignissen dazu führt, daß der Kandidat erhebliche Anteile der begrenzt zur Verfügung stehenden Zeit vergeudet und damit gar nicht mehr dazu kommt, in das Prüfungsgeschehen eingreifen zu können. Freilich ist es auch möglich, daß unsere aus der systemtheoretischen Analyse des Prüfungsgeschehens inhaltlich abgeleitete 'Kontrollstrategie' sich deswegen als Irrweg erweist, weil das Prüfungssystem,

das wir untersucht haben, den in der bisherigen Literatur konzipierten Leitlinien nicht folgt.

Eine Hilfestellung bei dieser Frage erhalten wir aus dem dritten Schritt der Analyse der Strategien, ihrer Einbettung in das Hypothesennetz.

Auf den ersten Blick ähnelt dieses Vorgehen dem der Konstruktvalidierung. In der Tat könnten wir diesen Anspruch erheben, wenn wir unsere Ergebnisse kreuzvalidiert hätten; wegen des erheblichen organisatorischen und rechentechnischen Aufwandes haben wir jedoch auf dieses Vorgehen verzichtet. Zudem haftet unseren Ergebnissen wegen der mit der Verwendung der KFA als Hauptanalysemethode einhergehenden Probleme ohnehin eine gewisse Heuristik an, die sich auch durch eine Kreuzvalidierung nicht aufheben ließe, sondern längere Zeiträume zur Überprüfung verlangt. Immerhin wäre es unredlich, wenn wir nicht auf die Gefahren eines Zirkelschlusses, die in unserer Vorgehensweise begründet ist, nicht hinweisen würden. Eine Kreuzvalidierung könnte wenigstens einige Zweifel ausräumen.

Tatsächlich entspricht unsere Vorgehensweise einer Mischung aus Hypothesengenerierung und Hypothesenprüfung, die wir - wären wir unserem Hintergrundinteresse der Prävention folgen können - unseres Erachtens hätten sukzessive anwenden müssen. Die angerissenen Probleme entstehen dadurch, daß wir diese verschiedenen Schritte in eine Arbeit integrieren.

Unsere Hypothese 4 lautete, daß erfolgreiche Kandidaten Kontrollstrategien realisieren, nicht-erfolgreiche die Strategie der Gelernten Hilflosigkeit.

Kernstrategieitems und Leistungsdaten sind nicht assoziiert; allenfalls kann man sagen, daß 'Assertive' ausreichend Informationen aus der Prüfung entnehmen, um ihre tatsächliche Zensur gut vorherzusagen zu können. Eindeutige Stellungnahmen zu Selbstverständlichkeiten gehen mit überdurchschnittlichen Leistungsdaten einher, 'differenzierte' Stellungnahmen mit unterdurchschnittlichen Leistungsdaten (nur bei VDN). Mäßig angespanntes-variabel-diversifizierendes Vorgehen geht (nur bei VDN) mit unterdurchschnittlichen Leistungsdaten einher.

Eindeutige Ergebnisse erhalten wir nur für den Satz der Trivialitems. 'Konforme' Antworten auf Selbstverständlichkeiten führen zu überdurchschnittlichen Ausprägungen, 'differenzierte' zu unterdurchschnittlichen Ausprägungen der Leistungsdaten. In Weiterführung dieser Ergebnisse zeigt sich, daß auch für den Datensatz des TVD zumindest der Teil der Hypothesen, der sich auf die Klasse der VDN bezieht, angenommen werden kann. Somit errechnen wir konvergente Ergebnisse, die für eine Assoziation der Merkmalsausprägung VD (der Variablen Prüfungsart) unterdurchschnittliche Leistungsdaten, mäßig angespanntes-variabel-diversifizierendes Vorgehen und 'differenzierte' Stellungnahmen zu den Trivialitems sprechen.

Gleichzeitig gelingt es nicht, konvergente Ergebnisse für eine Assoziation der entgegengesetzten Merkmals-

ausprägungen oben genannter Datensätze zu erhalten. Dies kann darauf zurückzuführen sein, daß die entsprechenden Konfigurationen durch die Hinzufügung der Variablen Prüfungsart 'aufgespalten' werden und als Folge größtenteils die Signifikanzgrenze verfehlen. Eine Hilfestellung liefert die Diskussion der Hypothese 5.

In unserer fünften Hypothese hatten wir angenommen, daß Kandidaten in Zwischenprüfungen häufiger die Strategie der Gelernten Hilflosigkeit realisieren als Kandidaten in Abschlußprüfungen, welche eher Kontrollstrategien realisieren.

Mit Ausnahme der Ergebnisse der Kernstrategieitems erhalten wir in den einzelnen Datensätzen nach Hinzufügen der Variablen Prüfungsart strukturell gleiche Ergebnisse: die Kollektive der 'Differenzierten' und der 'mäßig angespannt-variabel-diversifizierend Vorgehenden' werden der Merkmalsausprägung VD zugeordnet, während die jeweiligen Gegentypen 'aufgespalten' werden und nunmehr z. T. die Signifikanzgrenze verfehlen. Nach dem strengen Algorithmus annehmen können wir die Hypothese lediglich für den aus TVD(W) und TVD(H) zusammengesetzten Datensatz. Aufgrund der Ergebnisse der Kernstrategieitems müssen wir die in der Hypothese 5 enthaltene Kontrastbehauptung für diesen Datensatz explizit ablehnen.

Unsere siebte und letzte zu diskutierende Hypothese lautete, daß Kandidaten, die die Strategie der Gelernten Hilflosigkeit realisieren, stärkeren Prüfungsbelastungen unterworfen sind als Kandidaten, die die Kontrollstrategie einsetzen.

Wenn wir die Kernstrategieitems als Strategien verrechnen, müssen wir diese Hypothese ablehnen.

Verrechnen wir hingegen die Trivialitems als Strategien, so finden wir nicht nur die vorhergesagten Assoziationen, sondern auch die implizite Kontrastaussage in der Agglutinations-KFA mit den Analogskalen zur Erfassung der Postprüfungsbelastung voll bestätigt: 'Konforme' sind unterdurchschnittlich belastet, 'Differenzierte' skalieren überdurchschnittliche Merkmalsausprägungen. In den Agglutinations-KFAn mit den Trivialitems einerseits und den beiden anderen Belastungs-Datensätzen zeigen sich jeweils Assoziationen zwischen 'differenzierten' Strategien und überdurchschnittlichen Belastungen; das Kollektiv der 'Konformen' verfehlt jeweils die Signifikanzgrenze. Wiederholungsläufe für alle drei Agglutinations-KFAn unter Hinzufügung der Variablen Prüfungsart ergeben, daß in allen drei KFAn sich die Assoziationen 'differenzierter' Strategien mit überdurchschnittlichen Belastungen und dem Merkmal VD als signifikant erweist.

Wenn wir den dritten Datensatz zur Erfassung von Strategien (die beiden Versionen des TVD) verrechnen, so erhalten wir Ergebnisse, in denen stetiges und unfizierendes Vorgehen zu unterdurchschnittlichen Belastungen und der Kontrast variables und diversifizierendes Vorgehen zu überdurchschnittlichen Belastungen in Beziehung gesetzt wird. Dies entspricht den in Hypothese 7 postulierten Beziehungen. Aufgrund der Ergebnisse der Agglutinations-KFAn mit den Testverhaltensdifferentialen und den

Analogskalen zur Erfassung der Post-Prüfungs-Belastung einerseits und den Faktorscores der Beschwerdenliste andererseits müssen wir die Hypothese jedoch ablehnen.

Wir wollen die Ergebnisse derjenigen Hypothesen im Kontext betrachten, die den Begriff der Strategie als explikatives Konstrukt enthielten.

Allein die Trivialitems erbringen Ergebnisse, die die LH-theoretisch abgeleiteten Beziehungen (der einzelnen Datensätze untereinander und der Datensätze zu den Anthac-Klassen) abbilden.

Somit ergibt sich folgende Situation:

- Die Ergebnisse der parallelen Verhaltensstichproben lassen den Schluß zu, daß wir uns auf valide Messungen stützen.
- Alle Hypothesen, die das Konzept der Strategie nicht enthalten, können angenommen werden und führen - im Nachhinein - in gleicher Weise zu differenzierteren Aussagen, die dem Immunisierungsparadigma entsprechen.
- Eine A-priori-Operationalisierung von "Kontrollstrategie" und "Gelernter Hilflosigkeit" wurde nicht vorgenommen. Daher kann es - streng genommen - im Rahmen dieser Arbeit lediglich um den Nachweis gehen, daß es polare Strategien gibt. Dieser Nachweis ist formal ausgelegt, nämlich 1. durch den Nachweis von Gegentypen in den 3 Datensätzen und 2. durch deren Integration in die Hypothesennetze.

Dem zweiten formalen Kriterium genügen die Datensätze der Kernstrategieitems und der TVD-Scores nicht, wohl aber die Trivialitems.

- Im Nachhinein können wir nunmehr diskutieren, ob die gefundenen Zusammenhänge der Theorie der Gelernten Hilflosigkeit entsprechen und ob eine Übertragung dieser Theorie auf das Gebiet der Prüfungsbelastung möglich ist.

Eine Übertragung scheint uns möglich zu sein, wenn wir die in Kapitel 2.2.2.3.1 (S. 96 ff) dargestellte Erweiterung der LH-Theorie durch die Integration des Konfliktkonzepts aufgreifen: Dort haben wir dargestellt, daß die Unkontrollierbarkeitsbedingung auch dann erfüllt sein kann, wenn vorliegende diskriminative Reize von der Vp nicht kalkuliert werden. Wenn wir die Situation der Prüfung als diskriminativen Reiz betrachten, dann können wir vermuten, daß für Kandidaten, die sich nicht sicher sind, ob in dieser Situation bestimmte Ereignisse eintreten oder bestimmte Reaktionen absolut notwendig sind, die Unkontrollierbarkeitsbedingung erfüllt ist.

Die Gelernte Hilflosigkeit von Prüfungskandidaten zeigt sich in nicht-eindeutigen Reaktionen auf triviale Items; derjenige mit der Strategie der Gelernten Hilflosigkeit ist sich nicht völlig sicher, daß er zu Beginn der Prüfung aufgeregter ist als am Schluß, er ist nicht völlig sicher, daß es darauf ankommt, Zusammenhänge und Unterschiede zu sehen, zu beschreiben und anschaulich darzustellen. Ein Prüfungskandidat mit der Strategie der Gelernten Hilflosigkeit gestaltet damit eine - für ihn - fatale Situation. Wir wollen dies am Beispiel der drei konkreten Trivialitems darstellen.

Der Angst- und Aufregungsverlauf in mündlichen Prüfungen - darauf hatten wir bereits verwiesen - ist eine der wenigen gesicherten Erkenntnisse der Prüfungsforschung. Im theoretischen Teil haben wir ferner ausführlich dargestellt, daß es in mündlichen Prüfungen in erster Linie um kommunikative Operationen geht.

Die Kandidaten, denen wir die Strategie der Gelernten Hilflosigkeit zuordnen, sind demnach nicht sicher im Hinblick auf Ereignisse, die mit Sicherheit eintreten, und im Hinblick auf Operationen, die mit Sicherheit erforderlich sind.

Ein Kandidat mit der Strategie der Gelernten Hilflosigkeit muß unter der Bedingung begrenzt zur Verfügung stehender Zeit Ereignisse und Operationen kalkulieren, für die ein Kandidat mit Kontrollstrategie keine Sekunde investiert. Ein Kandidat mit der Strategie der Gelernten Hilflosigkeit nutzt also die ihm (begrenzt) zur Verfügung stehende Zeit nicht optimal; diese Schlußfolgerung steht im Einklang mit anderen - eingegrenzten - Prüfungsangstkonzepten (vgl. WINE, 1971; HECKHAUSEN, 1975) und dem Depressionskonzept von PAYNE & HEWLETT (1960).

Unsere drei Trivialitems dürften aber die Strategie der Gelernten Hilflosigkeit und die Kontrollstrategie nicht in toto erfassen, sondern stellen mit ziemlicher Sicherheit nur exemplarische Itemstichproben dar.

Von hierher können wir nun weitere notwendige Grundlagenforschung ableiten und auf präventive Ansätze zur Prüfungsbelastung eingehen.

Im theoretischen Teil haben wir abgeleitet, daß die Strategie der Gelernten Hilflosigkeit in die Operations-Suche beim Problemlösen eingreift, und haben drei Möglichkeiten genannt: die Suche wird vorschnell beendet (motivationales Defizit oder Einengung des Suchraums), die Suche wird auf irrelevante Operationen ausgedehnt (Ausweitung des Suchraums), und die Beurteilung geeigneter Operationen verzögert sich. Letztere beiden Möglichkeiten entsprechen mehr dem kognitiven Defizit. Von unseren Ergebnissen her können wir nun vermuten, daß eine Einengung des Suchraums eher für eine Kontrollstrategie steht und die Strategie der Gelernten Hilflosigkeit eher für eine Ausweitung des Suchraums und für unsichere Beurteilung der erforderlichen Operationen steht. Diese Hypothese nun können wir unter kontrollierten Laborbedingungen überprüfen.

Ebenso können wir direkte Schlußfolgerungen im Hinblick auf unser Hintergrundinteresse, der Prävention der Prüfungsbelastung, ziehen. Wenn Prüfungsbelastungen und schlechte Leistungen von Unsicherheit im Hinblick auf Selbstverständlichkeiten abhängt, dann besteht eine präventive Intervention in der Beseitigung dieser Unsicherheiten. Dies kann durch Aufklärungsaktionen unabhängig vom Prüfungssystem geschehen, aber auch durch (kleinere) Veränderungen des Prüfungssystems selbst.

In dieser Arbeit sind wir davon ausgegangen, daß Gelernte Hilflosigkeit als Strategie vom Prüfling in das System Prüfung eingebracht wird. Das Prüfungssystem selbst haben wir als konstant angenommen. Wir haben vernachlässigt, daß Prüfer selbst z. B. durch Erhöhung der Transparenz der Prüfungen dem Auftreten von Prüfungsbelastungen (und unterdurchschnittlichen Leistungen) vorbeugen können und umgekehrt, daß Prüfer durch 'Verschleiern' ihrer Prüfungen Prüfungsbelastungen (und unterdurchschnittliche Leistungen) induzieren können. Wir sind keineswegs der Auffassung, daß überdurchschnittliche Prüfungsbelastungen und unterdurchschnittliche Leistungen allein und ausschließlich eine Folge der Strategien der Kandidaten sind (dies ist vom theoretischen Hintergrund her nur eine Variante, die sich relativ leicht untersuchen ließ). Wir halten es durchaus für möglich, daß in einzelnen Prüfungsordnungen der Initiationsritus betont und Prüfungsbelastungen systematisch durch Verschleierungstechniken induziert werden. Im Gegensatz zu den im theoretischen Teil vorgestellten kulturanthropologischen Ansätzen und dem vorherrschenden Trend in den systemfunktionalistischen Konzepten gingen wir davon aus, daß Prüfungsbelastungen keine logische und notwendige Folge von Prüfungen schlechthin darstellen. Wir gingen davon aus, daß Prüfungsbelastungen dann auftreten, wenn Institution oder Prüfling dafür sorgen, daß Gelernte Hilflosigkeit als Strategie in die Prüfung eingebracht wird. Empirisch untersucht haben wir aus praktischen Gründen nur eine Variante. Immerhin ist das von uns vorgelegte Konzept der Kontrollstrategie ein Kontrapunkt im vorherrschenden Prüfungsforschungstrend.

Überraschend sind die Ergebnisse, die sich durch die Integration der Kernstrategien in das Hypothesennetz ergeben: das Kollektiv, das sein Verhalten als aktiv seine Meinung vertretend beschreibt, das sich ggf. gegen den Prüfer wehren will, mithin aktiv die Kommunikationssteuerung zu übernehmen bereit ist, fällt ebenso durch das Hypothesennetz wie der Gegentyp, der passiv bleibt und dem Prüfer die Initiative überlassen will, sein Verhalten zudem als submissiv beschreibt. Die theoretischen Gedankengänge, die eine aktiv kontrollierende Strategie in einen Zusammenhang mit der Prüfungssteuerung setzen - mit einer symmetrischen Prüfungsinteraktion - konnten nicht bestätigt werden. Zwar konnten wir nachweisen, daß es Kandidaten gibt, die diese Gedankengänge für sich aufgreifen, jedoch sind dies nicht die Kandidaten, deren Leistungs- und Belastungsdaten vorhersagbar wären.

Unsere Überlegungen standen ganz in der Tradition der bisherigen Prüfungsforschung, nach der eine mündliche Prüfung - überspitzt formuliert - ein Kampf zwischen zwei Gegnern darstellt, bei dem es für den Kandidaten darauf ankommt, sich als gleichwertig zu erweisen.

Tatsächlich scheinen wir jedoch auf ein mehr rationales Prüfungssystem gestoßen zu sein, in der der Aspekt des Initiationsritus endgültig aufgegeben wurde und in dem es tatsächlich nurmehr darauf ankommt, "daß den dazu eingesetzten Kommissionen Gelegenheit gegeben wird, sich zu überzeugen, daß man sich Überblick, Wissen und denkmäßige Schulung für den Beruf erworben hat!" (NEUMANN 1933, S. 48)

Unsere Ergebnisse lassen die Deutung zu, daß mit Beginn der Prüfung Unterschiede zwischen den Prüfern vernachlässigt werden können. Wir hielten es für plausibel, daß die Interaktion in Prüfungen deutlich von der Person des Prüfers abhängen müßte. Im Nachhinein stellen wir darum fest, daß unsere Eingrenzung der systemtheoretischen Analyse auf die Interaktion in der Prüfung zu eng war.

Die Unterbrechungen des Kandidaten durch den Prüfer stehen im Einklang mit dieser Schlußfolgerung. Wenn ein Prüfer einen Kandidaten unterbricht (gleichgültig, warum er dies tut), dann tut er dies Überzufällig häufig bei dem Kollektiv der 'Differenzierten' und bei denjenigen Kandidaten mit Überdurchschnittlichen Belastungssymptomen während der Prüfung. Die Unterbrechenshäufigkeit steht jedoch in keinem Zusammenhang mit den Leistungsdaten. Umgekehrt können wir nicht von 'differenzierten' Antworten oder starker Prüfungsbelastung auf die Unterbrechenshäufigkeit schließen. Unsere Ergebnisse lassen sich daher so interpretieren, daß 'differenzierte' Antworten und Überdurchschnittliche Belastungen die explizite Übernahme der Prüfungssteuerung durch den Prüfer zwar begünstigen, jedoch nicht notwendig nach sich ziehen. Möglicherweise sind Unterbrechungen durch den Prüfer ein Korrekturmechanismus im System, der schlechte Leistungsdaten verhindern soll: der Prüfer unterbricht den Kandidaten, bevor dieser ausreichend Indizien für eine unterdurchschnittliche Bewertung geliefert hat.

Weiterhin gibt es die Möglichkeit, daß Prüfer einen Kandidaten unterbrechen, weil dieser ein Thema so kompetent abhandelt, daß weitere Notenindizien nicht benötigt werden. In beiden Fällen stellen Unterbrechungen durch den Prüfer Aktionen dar, die nicht zuungunsten des Kandidaten initiiert werden.

Die von uns gefundenen Ergebnisse lassen die Transparenz von Prüfungen als zentralen Prädiktor für Leistungsniveau und Prüfungsbelastung interpretieren. Auf der psychologischen Seite zeigt sich dies als Kontrollstrategie (Transparenz), bzw. als Strategie der Gelernten Hilflosigkeit (mangelnde Transparenz).

Inwieweit unsere Ergebnisse eine Besonderheit des Bochumer Psychologischen Instituts oder gar die Verhältnisse zur Zeit der Datenerhebung widerspiegeln, müssen wir offen lassen. Ebenso bleibt offen, ob die Schlußfolgerungen für die Weiterentwicklung der Theorie der Gelernten Hilflosigkeit zwingend sind:

Das zentrale Element des LH-Effekts ist das kognitive Defizit, das aus mangelnder Kalkulation diskriminativer Reize besteht.

Unabhängig davon, daß aus methodischen Gründen die Ergebnisse unserer Arbeit kreuzvalidiert werden müssen (worauf wir aus organisatorischen und technischen Gründen verzichteten), müssen wir auf notwendige weitere Replikationen im Hinblick auf die Schlußfolgerungen verweisen: sowohl im Hinblick auf die Prüfungsaspekte wie im Hinblick auf die Theorie der Gelernten Hilflosigkeit.

5.3 Diskussion des hypothesengenerierenden Teils

Im Hypothesengenerierungskomplex A haben wir nach Hinweisen gefragt, die gegen die in dieser Arbeit vertretene Auffassung sprechen, die Person des Prüfers könne für die Analyse von Prüfungsbelastungen vernachlässigt werden. Anders formuliert: verhalten sich Prüfer wirklich wie Profis?

Unterschiede zwischen den Prüfern, die keinerlei Bezug zu Prüfungen haben, wollen wir ausklammern. Daß der eine Prüfer Zigarillos raucht, der andere seit kurzem auf Zigaretten verzichtet, daß der eine Prüfer einen Bart trägt, der andere stets glattrasiert erscheint, halten wir für irrelevant. In den bislang diskutierten Datensätzen haben wir lediglich bei den Leistungsdaten Unterschiede zwischen den Prüfern gefunden. Diese Unterschiede können aber auch auf prüfungsrelevante und lernintensive Lehrveranstaltungen in den zwei Semestern vor den Prüfungen zurückgeführt werden. In den Datensätzen zur Belastungsmessung haben wir keine univariaten Unterschiede gefunden; univariate Varianzanalysen ohne Berücksichtigung der Meßwiederholung sind aber relativ grobe Methoden, die vermutlich nur drastische Unterschiede wiedergeben.

Bei der Baseline-Datenerhebung haben wir die 'Angst vor' und die 'Anforderungen bei' den einzelnen Prüfern skalieren lassen. Die Ergebnisse sind mehrdeutig: Mit t-Tests (auf die wir wegen vieler Missing Values ausweichen mußten) lassen sich Unterschiede zwischen Prüfern sichern, ebenso finden wir in allen Fällen positive Korrelationen zwischen skalierten

Ängsten und Anforderungen; die Höhe der Korrelationen schwankt zwischen den einzelnen Prüfern jedoch in einem signifikanten Ausmaß; zudem ließen sich weder in der skalierten Angst noch mit der skalierten Anforderung Unterschiede zwischen Prüfern des gleichen Fachs sichern. Wir fanden demnach kein Indiz dafür, daß Angst oder Anforderungen bei der Auswahl der Prüfer eine nennenswerte Funktion zukommt. Wir fanden kein Indiz dafür, daß Prüfungskandidaten generell den Weg des geringsten Widerstandes gehen und sich immer für den angenehmsten und leichtesten Prüfer entscheiden.

Von den Beisitzern ließen wir die Häufigkeit skalieren, mit der ein Prüfer Emotionen sichtbar zeigt. Dabei zeigte sich, daß die Beobachterübereinstimmung (ermittelt mit Hilfe zufällig anwesender Prüfungszuschauer) ausgesprochen gering ist, daß insgesamt nur in einem Viertel der Prüfungen überhaupt irgendeine Emotion beobachtet wurde und daß wir keine überzufälligen Häufigkeiten für einen Prüfer oder durch einen Beisitzer fanden. Durch Analysen der einzelnen aufgeführten Emotionen gelangten wir zur Hypothese, daß - wenn überhaupt - ein Prüfer dann Freude zeigt. Die mangelnde Reliabilität der Skalen führen wir darauf zurück, daß die Prüfer - wenn überhaupt - dann nur kurzzeitig Emotionen zeigen. Zusammenfassend können wir demnach einen relativ neutralen Prüfungsstil konstatieren, der sich zudem gleichmäßig auf alle Prüfer verteilt. Die Feststellung, daß Prüfer sich wie 'Profis' verhalten und - wenn überhaupt - dann Emotionen der Freude zeigen, müßte allerdings in einer unabhängigen Studie kreuzvalidiert werden.

Im Hypothesengenerierungskomplex B haben wir die Frage aufgeworfen, ob sich Kandidaten in Zwischen- und Abschlußprüfungen im Hinblick auf die Prüfungsmeldemotivation unterscheiden. Vor allem für die Weiterentwicklung der Theorie der Gelernten Hilflo-sigkeit z. B. in Richtung auf die Theorie der Lei-stungsmotivation, ist es wichtig, zwischen kurzfristi-gen Verstärkern, die das Verhalten in den einzelnen Prüfungen steuern, und langfristigen Verstärkern zu unterscheiden.

Wir fanden einen Hinweis auf ein Kollektiv, das sich zur Prüfung anmeldet, weil die Angehörigen das er-warten, und das bei einem Mißerfolg (Nichtbestehen) soziale Unsicherheit verspüren würde, die Lästig-keit der erneuten Prüfungsvorbereitungen betont, insgesamt aber glaubt, über einen Mißerfolg hinweg-kommen zu können. Wir sehen hierin einen Hinweis auf soziale Motive, sich zur Prüfung anzumelden. Da dieser Typ nur auf dem 10-%-Niveau signifikant ist und seine Prägnanz auch nur gering ausfällt, müßte diese Konfiguration in einer Replikationsstudie erst kreuzvalidiert werden, bevor wir weitergehende Hy-pothesen generieren.

Weiterhin konnten wir zwei Kollektive nachweisen, die auf gegensätzliche Motive zur Prüfungsanmeldung bei VDN und HDN schließen lassen. Wir fanden, daß VDN in erster Linie in die Prüfungen gehen, um mit der Ausbildung weitermachen zu können, und als Gegen-satz, daß HDN in erster Linie aus finanziellen Grün-den in die Prüfung gehen und weil sie im Beruf etwas anderes machen wollen. Die eine Ausprägung der Ant-worten schließt die andere dabei jeweils aus.

Auf den ersten Blick scheinen diese Ergebnisse trivial zu sein, scheinen ein Spiegelbild der geltenden Prüfungsordnung darzustellen. Immerhin sind wir der erste Autor, der dieses akademische On-dit empirisch nachweist. Auf den zweiten Blick erkennen wir aber Möglichkeiten, die die Motivstruktur komplizieren können: VDn kalkulieren ausschließlich Inhalte, die ein Weiterstudium betreffen; wir wollen einmal unterstellen, sie kalkulieren Belohnungen. HDn geben an, etwas anderes machen und in den Beruf zu wollen; auch hier wollen wir unterstellen, es ginge um die Kalkulation von Belohnungen; gleichzeitig kalkulieren HDn jedoch finanzielle Aspekte, die wir als aversiv und somit als negative Verstärker auffassen können. Von daher ergibt sich die Hypothese, daß HDn im Vergleich zu VDn komplexere Motive haben, die zur Prüfungsanmeldung führen.

Wir haben diesen Aspekt nicht weiter verfolgt, da er nur die von uns getroffene Klassifikation der Kandidaten erfaßt. Wir haben keine gesicherten Ergebnisse, die über die Unterscheidung von Kandidaten in Zwischen- und von Kandidaten in Abschlußprüfungen hinausweisen.

Ebenfalls interessant in motivtheoretischem Sinne - und daher für die Weiterentwicklung der LH-Theorie wie der Prüfungsforschung relevant - sind die Ergebnisse der emotionalen Reaktion auf die Zensur und deren Kausalattribution. Die Ergebnisse, die wir bei der Auswertung der emotionalen Reaktion auf die Zensur erhielten, zeigen zunächst, daß die vorherrschenden Emotionen die der Erleichterung und der Freude sind. Dies muß sicherlich auf die insgesamt

'kandidatenfreundliche' Zensurierung als Folge des hohen Leistungsstandes zurückgeführt werden; in mehr als der Hälfte aller Prüfungen erfolgte die Benotung mit 'sehr gut' oder 'gut'. In der Agglutinations-KFA zeigt sich dann auch, daß überdurchschnittliche Leistungsdaten mit Freude, Erleichterung, Zufriedenheit, Zuversicht und Sicherheitsgefühlen einhergehen, unterdurchschnittliche Leistungsdaten mit den Gefühlen des Ärgers und der Enttäuschung. Dabei erweist sich die tatsächlich erhaltene Zensur als die bestimmende Variable. Unterdurchschnittliche Zensuren lösen demnach Gefühle aus, wie man sie nach einer Bestrafung oder der Verweigerung einer Belohnung erwarten könnte, überdurchschnittliche Zensuren stehen im Zusammenhang mit Gefühlen, die eine Belohnung und gleichzeitig negative Verstärkung signalisieren.

Die Analysen der Kausalattribution brachten insgesamt überraschende Ergebnisse. Zunächst einmal fanden wir Attributionsmuster, die einzeln gesehen mehr einer Erfolgsattribution entsprechen (was bei diesen Zensuren auch nicht weiter verwundert). Dann jedoch stießen wir auf ein vierdimensionales Attributionsystem, und es zeigte sich, daß überdurchschnittliche Zensuren zur simultanen Attribution auf Glück, Anstrengung, Fähigkeit und Prüfungssystem führen und daß unterdurchschnittliche Zensuren im Zusammenhang mit der Attribution auf Pech und Faulheit und Unfähigkeit und schlechtes Prüfungssystem stehen.

Betrachtet man unsere Ergebnisse noch genauer, so ergibt sich ein noch komplexeres Bild, auf das wir hier im einzelnen nicht eingehen wollen. Insgesamt

entsprechen unsere Ergebnisse nicht dem von WEINER et al. (1971) und anderen vorgelegten zweidimensionalen Modell. Dies mag mehrere Gründe haben, und jeder einzelne verdient es, in weiterer Forschung untersucht zu werden.

Zum einen kann es sein, daß Prüfungskandidaten nach der Prüfung in einer Stimmung sind, in der sie 'die ganze Welt umarmen' oder 'alles zum Teufel wünschen' wollen. Entsprechend undifferenziert müßte die Kausalattribution ausfallen. Eine andere Möglichkeit sehen wir darin, daß ein Modell, das für die Kausalattribution von Ringwurfspielen o. ä. überprüft wurde, nicht auf die Kausalattribution existentieller Leistungen übertragen werden kann, sondern hierfür ein komplexeres Modell benötigt wird, das Attributionen ermöglicht wie 'ich bin intelligent, weil ich mich (schon lange vorher) angestrengt habe; außerdem hatte ich auch noch Glück, gerade diesen Prüfer an diesem Tag erwischte zu haben' oder 'an meiner Intelligenz besteht kein Zweifel, aber bei diesem Prüfer muß man schon Glück haben (ich hatte Pech), um etwas zu erreichen'. Eine weitere Möglichkeit - und diese führt uns zur Theorie der Gelernten Hilflosigkeit - besteht darin, daß die Klassifikation von Ursachenerklärungen ein Irrweg ist und daß es wichtiger ist, inwieweit ein Prüfling zwischen der einen und der anderen Kausalattribution zu wechseln in der Lage ist. Dies würde darauf hinauslaufen, die Kausalattribution als Prozeß aufzufassen, der im Falle der retrospektiven Kausalattribution einen Anfang hat (das Ereignis, dessen Ursache zur Debatte steht) und ein Ende, das durch ein Ereignis oder durch eine 'Stopp-Regel' im Sinne von MILLER, GALANTER

& PRIBRAM (1960, deutsch: 1973) determiniert ist. Von diesem Modell bieten sich direkte Beziehungen zum Modell der Gelernten Hilflosigkeit als Strategie. Die Kontrollstrategie als Einengung des Suchraums für Operationen wäre dann als Einengung des Suchraums für Kausalattributionen aufzufassen, und die Gelernte Hilflosigkeit könnte dann als endloser Prozeß ohne eindeutiges Ergebnis aufgefaßt werden. Für das präventive Hintergrundinteresse käme es dann darauf an, den Kausalattributionsprozess zu beenden und ggf. in eine 'Schutz'-Attribuierung zu kanalisieren. Dieser Gedanke müßte jedoch erst in der Grundlagenforschung entfaltet und überprüft werden, was wiederum einen Schritt der Integration von angewandter klinisch-psychologischer und theoretischer Grundlagenforschung bedeuten würde.

6 ZUSAMMENFASSUNG

Die vorliegende Arbeit entstand bei dem Versuch, das Phänomen der 'Prüfungsangst' als klinisch-psychologisches Massenproblem zu analysieren und präventive Interventionsmöglichkeiten auszuloten. Eine differenzierte Version der Theorie der Gelernten Hilflosigkeit wurde auf das Phänomen 'Prüfungsangst' übertragen.

Prüfungskandidaten (in mündlichen Prüfungen) und Versuchspersonen in Experimenten (zur Gelernten Hilflosigkeit) stehen vor dem gleichen Problem: sie sind Interaktionspartner in einem dyadischen System und müssen eine gegebene Ausgangssituation (Prüfung, experimentelles Treatment) und einen gegebenen Zielzustand (eine möglichst gute Zensur, richtige Antworten) durch kommunikative Operationen miteinander verknüpfen und ineinander überführen. Bei der Suche nach geeigneten Operationen setzen Prüfungskandidaten und Versuchspersonen Strategien ein. In dieser Arbeit wird der Tatsache besondere Aufmerksamkeit gewidmet, daß Prüfungskandidaten (ebenso wie Versuchspersonen) durch die Anwendung von Strategien das Interaktionssystem mitgestalten, was sich nicht nur auf die emotionalen Belastungen, sondern auch auf die Leistungsdaten auswirkt. Der Deskription des Konstrukts 'Prüfungsstrategie' und seiner Stellung in einem Hypothesen- und Ergebnisnetz dienen wesentliche Teile dieser Arbeit.

Die Arbeit beschränkt sich auf die Analyse von

mündlichen Prüfungen von Studenten. Prüfungen sind ein notwendiger Bestandteil der Institution Universität und haben vier unterscheidbare Funktionen: sie steuern das Studium, sie dienen der Selektion, sie weisen einen Status zu und sie bilden den Abschluß der Hochschulsozialisation (bzw. eines Teils davon). Die wichtigsten Ansätze der bisherigen Prüfungsforschung werden vorgestellt; in den weitaus meisten wird die Universität als totale und unveränderbare Institution und der Prüfling als ein Ohnmächtig-Ausgelieferter konzipiert. Dieser vorherrschende Interpretationstrend zeigt sich auch in den wichtigsten Ansätzen der bisherigen Prüfungsangstforschung (die hier im einzelnen nicht diskutiert werden), in der aber eben diese Interpretation zu zahlreichen empirischen Arbeiten und somit zum pauschalen Nachweis führte, daß die Totalität der Institution mit der Symptomatik der Prüfungsangst assoziiert ist. Die Symptomatik der 'Prüfungsangst' besteht aus einem Angst-/Depressionsgemisch mit deutlichen Komponenten der Passivität, des Mißtrauens und des Pessimismus'. Diese Symptomatik ist mit schlechten Zensuren assoziiert.

In Abgrenzung zur vorherrschenden Interpretation wird in dieser Arbeit die Auffassung vertreten, daß Prüfungsbelastungen nicht zwingend notwendig allein als Folge der Situation Prüfung auftreten; vielmehr zwingt die Prüfung den Kandidaten zur Anwendung von Strategien, von denen eine - die Strategie der Gelerten Hilflosigkeit - Prüfungsbelastungen zur Folge hat.

Die Theorie der Gelernten Hilflosigkeit wird ausführlich und kritisch dargestellt, und es wird auf einige Weiterentwicklungsnotwendigkeiten hingewiesen. Die Theorie besagt, daß ein kognitives Element erlernt werden kann (das Element der Gelernten Hilflosigkeit), wonach es bei einem gegebenen Problem für die Verknüpfung von Ausgangssituationen und Zielzustand keine geeigneten Operationen gibt. Dieses kognitive Element wird als Strategie bei der Lösung anderer Probleme eingesetzt.

Dabei treten ein motivationales, ein kognitives und ein emotionales Defizit auf, die sich als verringerte Frequenz operanten Verhaltens, als Unfähigkeit, von Rückmeldungen zu profitieren, und als Angst-/Depressionsgemisch darstellen. Die Theorie macht Aussagen zum Auftreten der drei Defizite ebenso wie zu Vorbeugungsmöglichkeiten. In der Forschung zur Gelernten Hilflosigkeit werden drei methodische Ansätze verwendet: die Klassifikation der Versuchstiere oder Versuchspersonen entsprechend ihrer Vorerfahrungen (die die Defizite begünstigen oder ihnen vorbeugen), die Realisation entsprechender Vorerfahrungen im Experiment und die (direkte oder indirekte) Messung des kognitiven Elements der Gelernten Hilflosigkeit.

In dieser Arbeit werden die Methodenvarianten der Klassifikation und der direkten und indirekten Messung aufgegriffen.

Prüfungskandidaten werden danach klassifiziert, ob sie Zwischen- oder Abschlußprüfungen ablegen. Für

die Messung der Strategien werden mehrere Variablen-sätze eingesetzt. Einer entspricht dem Hypothesennetz.

Die Ergebnisse eines natürlichen Experiments mit mehrfach wiederholter Messung zeigen, daß die Prüfungsstrategie der Gelernten Hilflosigkeit durch einen Mangel an diskriminativen Reizen gekennzeichnet ist, der sowohl zu unterdurchschnittlichem Abschneiden als auch zu überdurchschnittlichen Belastungen führt, zudem in erster Linie Kandidaten in Zwischenprüfungen betrifft. Der Gegensatz dieser Strategie - die Kontrollstrategie - kann weniger deutlich in einen Zusammenhang mit überdurchschnittlichem Abschneiden und unterdurchschnittlichen Belastungen gesetzt werden. Die Strategie der Gelernten Hilflosigkeit stellt sich beim Problemlösen als Ausweitung des Suchraums für geeignete Operationen dar, was beim Prüfungskandidaten zu schlechtem Ausnützen der begrenzten Zeit führt. Die Kontrollstrategie besteht in einer Einengung des Suchraums (mit optimaler Nutzung der begrenzten Zeit).

Hauptanalysemethode ist die Konfigurationsfrequenzanalyse. Ihre Vor- und Nachteile werden ausführlich diskutiert, einige heuristische Interpretationsschritte werden dargestellt und einige Beispiele demonstrieren den Unterschied zu anderen statistischen Verfahren.

In der Diskussion werden weitere notwendige Grundlagenforschungen erörtert. Zudem wird das ursprüngliche Erkenntnisinteresse wieder aufgegriffen: Prü-

funksbelastungsprophylaxe besteht demnach in erster Linie in der Erhöhung der Transparenz von Prüfungen, was aber - als Nebeneffekt - gleichzeitig zu einem Ansteigen der Leistungsindikatoren führen müßte.

7 ANHANG

Seite

7.1 Fragebogen "Prüfungsstress" (postalischer Versand ca. 14 Tage vor der ersten mündlichen Prüfung)	468
7.2 Fragebogen AN (Bearbeitung <u>nach</u> der Prüfung <u>vor</u> Bekanntgabe der Zensur)	499
7.3 Fragebogen F2 (Bearbeitung <u>nach</u> Bekanntgabe der Zensur)	503
7.4 Fragebogen F3 (Bearbeitung <u>nach</u> Bekanntgabe der Zensur und nach F2)	515
7.5 Listen, Schreiben und Beisitzerblatt	522
7.6 Auswertungsschlüssel und Programme	528
7.7 Liste der in den Konfigurationsfrequenzanalysen verrechneten Variablen, Angaben zur Klassenbildung und zur Interpretation der Klassen	534
7.8 Mittelwerte, Standardabweichungen und Missing Values der Items der Strategiensammlung getrennt nach Datenerhebungszeitpunkten	537

7.1 Fragebogen "Prüfungsschritt"

postalischer Versand am 16.01.76, ca. 14 Tage vor
der ersten mündlichen Prüfung

Der hier wiedergegebene Fragebogen gliedert sich
wie folgt:

	Seite
2 Seiten Erläuterungen für den Kandidaten	470-471
1 Seite <u>Fragen zur Person</u>	472
1 Seite <u>Globaleinschätzungen</u> von Prüfern und Personen	473
2 Seiten Fragen zu <u>Vorerfahrungen</u> mit münd- lichen Prüfungen	474 475
2 Seiten Fragen zur Prüfungsanmeldemotiva- tion (" <u>Konsequenzen</u> ")	476 477
1 Seite Attribuierungsskalen zum <u>besser</u> und <u>schlechter</u> als erwarteten Ab- schneiden	478

Prüfungsvorbereitungen

4 Seiten Studiengewohnheitsinventar (SGI mit 1 Seite Erläuterung und 50 Items)	479-482
1 Seite Testverhaltensdifferential (TVD- <u>Habe</u>)	483

	<u>Emotionstests</u>	Seite
1 Seite	Erläuterungen zu den Emotionsskalen	484
1 Seite	Analogskalen	485
1 Seite	Hilflosigkeitsskala (HiS)	486
1 Seite	Hoffnungslosigkeitsskala (HoS)	487
1 Seite	BECK-Depressions-Inventar, Selbst- Rating Kurzform (BDI)	488
1 Seite	State-Trait-Anxiety-Inventory, Form X-1 (STAI)	489
4 Seiten	<u>Freiburger Beschwerdeliste</u> (FBL)	490-493

	<u>Prüfungsstrategien</u>	
3 Seiten	Strategiensammlung (SS) mit 34 Items	494-496
1 Seite	Testverhaltensdifferential (TVD-Werde)	497
1 Seite	Dank an die Prüflinge	498

Für die versandten Fragebögen wurden die Teile

1. Prüfungsvorbereitungen (SGI und TVD-H),
2. Emotionstests (Analogskalen, HiS, STAI, BDI, HoS),
3. FBL und
4. Prüfungsstrategien (SS und TVD-W)

nach dem Schema der Tabelle A und in dieser Reihenfolge permutiert. Innerhalb der Emotionstests wurde nach dem Schema der Tabelle B und in dieser Reihenfolge permutiert.

Tabelle A

1234
1243
1324
1342
1423
1432

2134
2143
2314
2341
2413
2431

3124
3142
3214
3241
3412
3421

4321
4312
4213
4231
4123
4132

Permutationen
der Fragebogen-
teile

Tabelle B

12345
31245
34512
12453
45123
45312

54321
54213
21543
35421
32154
21354

31425
31254
43125
42531
25431
25314

52431
45213
52134
13524
13452
41352

23541
23415
52341
54123
41523
41235

14532
51423
14325
32145
32514
53214

34512
23451
45123
51234
42135
32154

Permutationen
der Emotions-
tests

RUHR-UNIVERSITÄT BOCHUM

PSYCHOLOGISCHES INSTITUT
AE Klinische Psychologie
Lothar Breitkopf

463 BOCHUM-QUERENBURG, 14.1.76
UNIVERSITÄTSSTRASSE 150
GEBÄUDE GAFO 04/376
POSTFACH 2148
FERNRUF (02 34) ~~71~~ 700/4610
TELEX 08 25 860

Projekt »Prüfungsstress«

Liebe(r) Prüfungskandidat(in) des Frühjahrstermins!

Wie Du Dich vielleicht noch von der Verteilung der Prüfungstermine her erinnerst, führe ich im Zusammenhang mit den Frühjahrsprüfungen eine Untersuchung über die Belastungen von Prüfungskandidaten durch. Ich habe Dich gebeten, an der Untersuchung als Versuchsperson mitzuwirken, und bitte Dich jetzt, diesen Fragebogen auszufüllen.

Er ist etwas länger geworden, als ich beabsichtigt habe. Die Dicke des Fragebogens täuscht jedoch: das Ausfüllen dürfte ca. 45 - 50 Minuten dauern.

Ich möchte Dich bitten, diesen Teil einer Stunde zur Mitarbeit zu investieren. Natürlich wirst Du Dich fragen, wozu das gut ist. Zum einen profitiere ich, weil ich für meine Dissertation über "Prophylaxe der Prüfungsangst" dieses empirische Material benötige. Zum anderen - so hoffe ich - profitieren vom Anwendungsaspekt meiner Dissertation zukünftige Examenskandidaten. Dir selber nützt eine Mitarbeit demnach nichts, Du würdest jedoch späteren Generationen von Prüflingen helfen.

Du würdest Dir aber auch nicht schaden! Ich habe nämlich Vorsorge getroffen - und werde das auch weiter tun -, daß aus den Ergebnissen der Dissertation keine Rückschlüsse auf einzelne Kandidaten möglich sein werden. Es ist dabei selbstverständlich, daß vor dem Ende der letzten Prüfung dieses Termins kein Ergebnis in irgendeiner Form bekannt wird. Kein Professor wird jemals erfahren, wie Du über ihn denkst! Die Erhebung der Daten geschieht anonym!

Während der Prüfungszeit werde ich mich noch zwei oder drei mal

an Dich wenden. Es geht darum, jeweils nach der Prüfung, jedoch vor Bekanntgabe der Zensur - in den 2 Minuten, die Du sowieso wartest -, einige Skalen anzukreuzen. Wenn Du die Zensur erfahren hast, wirst Du einen Fragebogen erhalten, dessen Ausfüllen ca. 20 Minuten dauern wird. Im Moment steht fest, daß dies die Prüfungen bei Heckhausen und bei Schulte betrifft (weil von denen alle Kandidaten geprüft werden). Den 3. Termin muß ich noch ausfüllen; er läge zwischen den beiden obigen Prüfungen.

Nun zu dem beigegeführten Fragebogen! Es ist sehr wichtig, daß Du alle Fragen und Items beantwortest. Es ist nicht nötig, daß Du stundenlang über eine einzelne Frage nachdenkst! Antworte so, wie es Dir im ersten Moment oder am ehesten plausibel erscheint! Gehe zügig, aber genau vor!

Sollte Dir etwas unklar sein (z.B. die Art der Antwort), dann beantworte die Frage so, wie es in Fragebögen üblich ist.

Wie Du Dir vielleicht vorstellen kannst, benötige ich für eine Mehrfacherhebung einen Personenidentifikationscode. Ein Code, den niemand vergißt, ist z.B. seine Matrikelnummer.

Trage sie bitte hier ein.

--	--	--	--	--

Es versteht sich von selbst, daß die Matrikelnummer von mir noch umcodiert wird, und die vor Dir liegende Seite dann vernichtet wird. Dadurch ist sichergestellt, daß niemand Deine Angaben identifizieren kann. Wenn Du immer noch Bedenken haben solltest, dann wähle einen anderen Code, z.B. Geburtsdatum einer Freundin oder die Quersumme dieser Zahl. Mache in diesem Falle bitte ein Kreuz neben die obigen Felder.

Wenn Du diesen Fragebogen erhalten hast, bearbeite ihn möglichst noch heute. Du hast folgende Möglichkeiten, ihn zurückzusenden:

- mit dem beigelegten, bereits frankierten Umschlag durch die Post,
- persönlich im Sekretariat Schulte (Linda Mühndorf), GAFO 04/916, daselbst in mein Fach, oder
- bei Frau Radermacher (Institutssekretariat) in das Fach Schulte/Klinische Psychologie.

ES IST WICHTIG, DASS JEDER MITMACHT (so viele Kandidaten gibt es nämlich nicht) !

BEARBEITE DIE FRAGEN ZÜGIG, ABER GENAU !

MATRIKELNUMMER BITTE NICHT VERGESSEN !

DIE ERHEBUNG IST ANONYM !

Globaleinschätzungen

- Wieviel Angst hast Du (bzw. hättest Du) vor einer Prüfung bei

	überhaupt keine	durch- schnittlich	so stark wie noch nie
Schulte	0---0---0---0---0---0---0---0---0---0		
Hörmann	0---0---0---0---0---0---0---0---0---0		
Heckhausen	0---0---0---0---0---0---0---0---0---0		
Lelius	0---0---0---0---0---0---0---0---0---0		
Schönbach	0---0---0---0---0---0---0---0---0---0		
Graefe	0---0---0---0---0---0---0---0---0---0		
Rutenfranz	0---0---0---0---0---0---0---0---0---0		?

- Wie schätzt Du die Prüfungsanforderungen ein bei

	sehr niedrig	durchschn.	extrem hoch
Heckhausen	0---0---0---0---0---0---0---0---0---0		
Lelius	0---0---0---0---0---0---0---0---0---0		
Schulte	0---0---0---0---0---0---0---0---0---0		
Hörmann	0---0---0---0---0---0---0---0---0---0		
Rutenfranz	0---0---0---0---0---0---0---0---0---0		
Graefe	0---0---0---0---0---0---0---0---0---0		
Schönbach	0---0---0---0---0---0---0---0---0---0		?

(Bitte jeden Prüfer einschätzen)

- Welche Zensur erwartest Du als Gesamtzensur?

0-eher 1, 0-eher 2, 0-eher 3, 0-eher 4

- Wie wären Deine Gefühle bei einer Gesamtzensur "3"?

Freude	Gleich- gültigkeit	starker Ärger
0---0---0---0---0---0---0---0---0		

- Wie wären Deine Gefühle bei einer Gesamtzensur "1"?

extreme Freude	völlige Verblüffung
0---0---0---0---0---0---0---0---0	

- Welche Zensur erwartest Du in der Mehrheit der mündlichen Einzelprüfungen?

0-eher 1, 0-eher 2, 0-eher 3, 0-eher 4

- Wieviel Tage sind es noch bis zu Deiner ersten mündlichen Prüfung?

--	--

- Welches Datum schreiben wir heute (der Tag, an dem Du diesen Fragebogen ausfüllst) ?

--	--

1. 1976

Vorerfahrungen

Welche Vorerfahrungen hast Du mit mündlichen Prüfungen?

Anzahl der an einer Universität bis jetzt abgelegten mündlichen Prüfungen.

--	--

Wieviel mündliche Prüfungen hast Du beim Abitur (oder Abitur-äquivalent) abgelegt?

--	--

Wieviele von den bisherigen mündlichen Prüfungen (Abitur und Uni) waren den bevorstehenden vergleichbar in

Arbeitsaufwand

--	--

Schwere

--	--

?

Glaubst Du, daß Du den Prüfern innerhalb der letzten 2 Jahre irgendwann aufgefallen bist durch (Zutreffendes bitte ankreuzen!)

- ☐ Beteiligungen in Seminardiskussionen,
- ☐ schriftliche Arbeiten,
- ☐ politische Ansichten,
- ☐ anderweitig, nämlich im Zusammenhang mit
- ☐ nein, glaube ich nicht ?

Hattest Du im Zusammenhang mit der Mehrheit der bisherigen mündlichen Prüfungen Angst? (Zutreffendes bitte ankreuzen!)

☐ trifft nicht zu

☐ bereits lange vorher

☐ unmittelbar vorher

☐ während der Prüfung

Bist Du im Zusammenhang mit einer Prüfung schon einmal krank-
geworden?

☐ ja

☐ nein

☐

vegetative Dystonie

☐

Erschöpfungsdepression

☐

..... (bitte ein-
tragen!)

Bist Du schon einmal durch eine mündliche Prüfung durchgefallen?

☐ ja

☐ nein



Kamst Du deswegen in finanzielle Schwierigkeiten?

☐ ja

☐ nein

Hat jemand Dich deswegen gehänselt?

☐ ja

☐ nein

Hast Du aus diesem Grund an Deinen Fähigkeiten gezweifelt?

☐ ja

☐ nein

Hast Du Dir deshalb einen Ausbildungsgang verbaut?

☐ ja

☐ nein

Konsequenzen

Warum hast Du Dich für die Prüfungen gemeldet?

(Bitte alles Zutreffende ankreuzen!)

- ☐ Weil man üblicherweise genau in meinem Semester die Prüfungen ablegt,
- ☐ weil ich in den Beruf will,
- ☐ aus finanziellen Gründen,
- ☐ damit ich mit meiner Ausbildung weitermachen kann,
- ☐ weil ich sowieso schon zu lange studiert habe,
- ☐ weil man mir eine Stelle angeboten hat,
- ☐ weil das bisherige Studium langweilig geworden ist,
- ☐ weil meine Angehörigen (Eltern, Freund/in, Frau o.a.) das erwarten,
- ☐ weil ich ausreichend vorbereitet bin,
- ☐ weil ich etwas anderes machen möchte.

Gesetzt den Fall, Du müßtest eine Prüfung wiederholen; was würde geschehen? (Bitte alles Zutreffende ankreuzen!)

- ☐ Meine Angehörigen wären ausgesprochen argerrlich.
- ☐ Ich käme in finanzielle Schwierigkeiten.
- ☐ Ich wäre vor lauter Scham sehr unsicher in Umgang mit anderen.
- ☐ Das würde wegen des erneuten Arbeitsaufwandes für mich die Grenze der Belastbarkeit darstellen.
- ☐ Das würde für mich wirklich nichts anderes bedeuten als ein verplumptes Semester.

- ☐ Ich würde zwar die Wiederholungsprüfung anstreben, wäre aber unsicher, ob ein Weiterstudium noch Zweck hat.
- ☐ Im ersten Augenblick würde ich die Welt nicht mehr verstehen; darüber würde ich aber hinwegkommen.
- ☐ Das ist eine unbeantwortbare Frage, weil die Wahrscheinlichkeit durchzufallen für mich nur theoretisch besteht.
- ☐ Das würde mich nicht wundern, weil ich für einige Prüfungen sowieso ein Risiko einkalkuliert habe.
- ☐ Das wäre sehr ärgerlich, würde mich aber nicht umwerfen, weil ich es ja eigentlich nicht verhindern konnte.

Gesetzt den Fall, Du würdest in einer Prüfung besser abschneiden, als Du selbst es erwartet hast. In welchem Ausmaß würdest Du die Gründe dafür in den folgenden Ursachen sehen?

	vermut-				
	bestimmt	lich	vielleicht	ziemlich	völlig
	nicht	nicht	doch	sicher	sicher
1. Ich habe mich angestrengt.	0-----	0-----	0-----	0-----	0-----
2. Ich hatte Glück.	0-----	0-----	0-----	0-----	0-----
3. Es liegt an meinen Fähigkeiten.	0-----	0-----	0-----	0-----	0-----
4. Der Prüfer prüft gut.	0-----	0-----	0-----	0-----	0-----
5. Ich war faul.	0-----	0-----	0-----	0-----	0-----
6. Ich hatte Pech.	0-----	0-----	0-----	0-----	0-----
7. Das liegt an meiner Unfähigkeit.	0-----	0-----	0-----	0-----	0-----
8. Der Prüfer prüft schlecht.	0-----	0-----	0-----	0-----	0-----
9. Ich war fleißig.	0-----	0-----	0-----	0-----	0-----
10. Das war Zufall.	0-----	0-----	0-----	0-----	0-----
11. Ich bin intelligent.	0-----	0-----	0-----	0-----	0-----
12. Es wird nicht viel verlangt.	0-----	0-----	0-----	0-----	0-----

Gesetzt den Fall, Du würdest in einer Prüfung schlechter abschneiden, als Du selbst erwartet hast. In welchem Ausmaß würdest Du die Gründe dafür in den folgenden Ursachen sehen?

	vermut-				
	bestimmt	lich	vielleicht	ziemlich	völlig
	nicht	nicht	doch	sicher	sicher
1. Ich habe mich angestrengt.	0-----	0-----	0-----	0-----	0-----
2. Ich hatte Glück.	0-----	0-----	0-----	0-----	0-----
3. Es liegt an meinen Fähigkeiten.	0-----	0-----	0-----	0-----	0-----
4. Der Prüfer prüft gut.	0-----	0-----	0-----	0-----	0-----
5. Ich war faul.	0-----	0-----	0-----	0-----	0-----
6. Ich hatte Pech.	0-----	0-----	0-----	0-----	0-----
7. Das liegt an meiner Unfähigkeit.	0-----	0-----	0-----	0-----	0-----
8. Der Prüfer prüft schlecht.	0-----	0-----	0-----	0-----	0-----
9. Ich war fleißig.	0-----	0-----	0-----	0-----	0-----
10. Das war Zufall.	0-----	0-----	0-----	0-----	0-----
11. Ich bin intelligent.	0-----	0-----	0-----	0-----	0-----
12. Es wird nicht viel verlangt.	0-----	0-----	0-----	0-----	0-----

Die folgenden Items beziehen sich auf Deine konkreten Studiengewohnheiten und konkreten Prüfungsvorbereitungen. Insgesamt 50 Behauptungen kannst Du für Dich entweder ablehnen oder annehmen. Die vorgegebenen Antwortkategorien lauten:

stimmt genau	stimmt in etwa	weiß nicht	eher im Gegenteil	ganz im Gegenteil
☒	☐	☐	☐	☐

Bei diesen Items bitte die entsprechende "O" ankreuzen!

Bearbeite bitte jedes Item!

Wähle diejenige Antwort, die Deinen Ansichten entspricht!

	stimmt genau	stimmt in etwa	weiß nicht	eher im Gegen- teil	ganz im Gegen- teil
1. Ich glaube, daß meine persönliche Methode der Vorlesungsverarbeitung wirkungsvoll ist und den Studienanforderungen genügt	0 ----	0 ----	0 ----	0 ----	0
2. Ich arbeite in letzter Zeit lieber mit Hintergrundmusik	0 ----	0 ----	0 ----	0 ----	0
3. Genau wie früher werde ich auch jetzt mit meiner Arbeit meist fertig	0 ----	0 ----	0 ----	0 ----	0
4. Es fällt mir schwer, mit der Arbeit anzufangen	0 ----	0 ----	0 ----	0 ----	0
5. Ich arbeite in letzter Zeit bis spät in die Nacht	0 ----	0 ----	0 ----	0 ----	0
6. Wenn ich Lehrbücher oder Zeitschriftenartikel lese, mache ich mir regelmäßig Notizen	0 ----	0 ----	0 ----	0 ----	0
7. Es kam häufiger vor, daß ich einen großen Teil einer Vorlesung in Gedanken bei etwas anderem verbrachte	0 ----	0 ----	0 ----	0 ----	0
8. Wenn ich mir in einer Vorlesung Notizen machte, habe ich meist so schnell wie möglich geschrieben	0 ----	0 ----	0 ----	0 ----	0
9. In Vorlesungen, Übungen und Seminaren bin ich zuweilen so sehr mit meinen Notizen beschäftigt, daß ich den Faden verliere	0 ----	0 ----	0 ----	0 ----	0
10. Ich habe einen festen Arbeitsplatz für meine Prüfungsvorbereitungen	0 ----	0 ----	0 ----	0 ----	0
11. Auf meinem Schreibtisch liegen Dinge, die mich zuweilen vom Lernen ablenken	0 ----	0 ----	0 ----	0 ----	0
12. Wenn ich bei einem englischen Text auf ein Wort stoße, das ich nicht kenne, schlage ich es im Wörterbuch nach	0 ----	0 ----	0 ----	0 ----	0
13. Bei einem Zeitschriftenartikel lese ich in jedem Fall erst die Zusammenfassung, bevor ich den Text durcharbeite	0 ----	0 ----	0 ----	0 ----	0
14. Bei den Prüfungsvorbereitungen versuche ich, so viel wie möglich auswendigzulernen	0 ----	0 ----	0 ----	0 ----	0
15. Nach wie vor lese ich regelmäßig eine Tageszeitung	0 ----	0 ----	0 ----	0 ----	0
16. In Seminaren diskutiere ich zuweilen eifrig mit	0 ----	0 ----	0 ----	0 ----	0

	stimmt genau	stimmt in etwa	weiß nicht	eher im Gegen- teil	ganz im Gegen- teil
17. Wenn ich Schwierigkeiten mit dem Stoff habe, frage ich um Rat	0	----	0	----	0
18. Vor dem Lesen überfliege ich ein Kapitel in einem Lehrbuch	0	----	0	----	0
19. Ich habe Mühe, mich schriftlich auszudrücken	0	----	0	----	0
20. Beim Lesen lasse ich Tabellen und graphische Darstellungen unbeachtet	0	----	0	----	0
21. Für jede Prüfung habe ich einen Extraort (-ordner, -mappe) für spezielle Notizen	0	----	0	----	0
22. Morgens vor Arbeitsbeginn frühstücke ich ausgiebig und gemütlich	0	----	0	----	0
23. Zur Zeit suche ich häufig die Institutsbibliothek auf	0	----	0	----	0
24. Persönliche Probleme oder Tagesereignisse belasten mich in letzter Zeit so stark, daß ich mich kaum auf die Arbeit konzentrieren kann	0	----	0	----	0
25. Ich habe meine festen Arbeitszeiten vorher festgelegt und halte mich i.A. an diesen Plan	0	----	0	----	0
26. Ich teile mir meine Arbeit häppchenweise ein, dennoch habe ich selten den Eindruck, vorwärts zu kommen	0	----	0	----	0
27. Der Stoff, den ich in diesen Tagen durcharbeite, ist mir zumindest in Umrissen meist schon bekannt	0	----	0	----	0
28. Manchmal merke ich, daß ich etwas wohl visuell gelesen, geistig aber nicht aufgenommen habe	0	----	0	----	0
29. Pausen zwischen Arbeitsblöcken kann ich mir in diesem Tagen nicht leisten	0	----	0	----	0
30. Sehr häufig könnte ich alle Bücher gegen die Wand werfen	0	----	0	----	0
31. Besonders in letzter Zeit habe ich den Eindruck, daß die Prüfungspflichtlektüre nur ganz langsam weniger wird	0	----	0	----	0
32. Ich bin mir völlig sicher, daß ich den Stoff für einige Prüfungen schaffe	0	----	0	----	0

	stimmt genau	stimmt in etwa	weiß nicht	eher im Gegen- teil	ganz im Gegen- teil
33. Ich mache alle 1 1/2 - 2 Stunden Arbeitspausen	0	----	0	----	0
34. In letzter Zeit werde ich von der Arbeit überrollt	0	----	0	----	0
35. Ich bereite mich auf die Prüfungen allein vor	0	----	0	----	0
36. Mein Gedächtnis läßt in letzter Zeit zu wünschen übrig	0	----	0	----	0
37. Manchmal überlege ich, ob ich mich auf die Prüfungen richtig vorbereite	0	----	0	----	0
38. Mir fällt es nicht schwer, 4 - 5 Stunden ununterbrochen zu lernen	0	----	0	----	0
39. Ich bin unsicher, ob ich auch nur eine Prüfung so weit vorbereiten kann, daß ich wenigstens die Pflichtlektüre gelesen habe	0	----	0	----	0
40. Meine Prüfungsvorbereitungen beinhalten Diskussionen in einer Arbeitsgruppe	0	----	0	----	0
41. Wenn ich in letzter Zeit merke, daß ich etwas nicht begriffen habe, werde ich sehr unsicher	0	----	0	----	0
42. Mir wäre es lieber, wenn mir jemand sagen würde, wie ich vorzugehen habe	0	----	0	----	0
43. Zuweilen macht mir das Lernen Spaß	0	----	0	----	0
44. Ich lasse mich bei den Prüfungsvorbereitungen abhören	0	----	0	----	0
45. Wenn ich höre, was andere Kandidaten schaffen, werde ich unsicher	0	----	0	----	0
46. Insgesamt glaube ich, mich intensiver als andere vorzubereiten	0	----	0	----	0
47. Ich werde durch die Schwarzmalerei anderer beunruhigt	0	----	0	----	0
48. Es fällt mir schwerer als anderen, den Stoff zu behalten	0	----	0	----	0
49. Neben dem, was ich sonst tue, bleibt mir kaum Zeit, mich auf die Prüfungen vorzubereiten	0	----	0	----	0
50. Ich bin der Ansicht, daß 14 Tage Vorbereitung pro Prüfung voll ausreichen	0	----	0	----	0

Im folgenden möchte ich Dich bitten, Deine bisherigen Prüfungsvorbereitungen einzuschätzen. Dazu sind 15 Aussagenalternativen aufgeführt. Zwischen den Alternativen ist eine siebenstufige Skala. Du kannst nun durch Ankreuzen angeben, ob Du eher so vorgegangen bist, wie es in der linken Alternative behauptet wird: dann mach bitte ein Kreuz auf der linken Seite der Skala. Trifft eher die rechte Alternative zu, dann mach Dein Kreuz bitte in der rechten Hälfte. Dabei kannst Du abstufen: wenn eine Alternative ganz genau zutrifft, dann kreuze bitte die "3" an; wenn die eine Alternative mehr zutrifft als die andere, dann kreuze bitte die "1" oder "2" an. Die "0" bedeutet: "weiß nicht" oder "kann mich nicht entscheiden". Kreuze die "0" bitte nur dann an, wenn für Dich eine Entscheidung völlig unmöglich ist.

Habe bis jetzt phantasievoll gedacht	3 2 1 0 1 2 3 0-0-0-0-0-0-0-0-0-0	Habe bis jetzt nüchtern gedacht
Bin bis jetzt besonnen vorgegangen	3 2 1 0 1 2 3 0-0-0-0-0-0-0-0-0-0	Bin bis jetzt zügig vorgegangen
Habe bis jetzt den Gegenstand ganzheitlich aufgefaßt	3 2 1 0 1 2 3 0-0-0-0-0-0-0-0-0-0	Habe bis jetzt den Gegenstand zergliedert
Habe bis jetzt die Zusammenhänge aufgelöst	3 2 1 0 1 2 3 0-0-0-0-0-0-0-0-0-0	Habe bis jetzt Zusammenhänge hergestellt
Bin bis jetzt leichtfüßig vorgegangen	3 2 1 0 1 2 3 0-0-0-0-0-0-0-0-0-0	Bin bis jetzt hartnäckig vorgegangen
Habe bis jetzt die Sache in den Griff genommen	3 2 1 0 1 2 3 0-0-0-0-0-0-0-0-0-0	Habe bis jetzt die Sache auf mich wirken lassen
Habe bis jetzt das Verhältnis der Einzelheiten zu einem übergeordneten Ganzen wichtig genommen	3 2 1 0 1 2 3 0-0-0-0-0-0-0-0-0-0	Habe bis jetzt das Verhältnis der Einzelheiten untereinander betont
Bin bis jetzt rasch vorgegangen	3 2 1 0 1 2 3 0-0-0-0-0-0-0-0-0-0	Bin bis jetzt gründlich vorgegangen
Bin bis jetzt systematisch vorgegangen	3 2 1 0 1 2 3 0-0-0-0-0-0-0-0-0-0	Bin bis jetzt ohne strenge Systematik vorgegangen
Bin bis jetzt variabel vorgegangen	3 2 1 0 1 2 3 0-0-0-0-0-0-0-0-0-0	Bin bis jetzt stetig vorgegangen
Habe mich bis jetzt angestrengt	3 2 1 0 1 2 3 0-0-0-0-0-0-0-0-0-0	Bin bis jetzt gelassen geblieben
Habe bis jetzt den Kern der Sache beachtet	3 2 1 0 1 2 3 0-0-0-0-0-0-0-0-0-0	Habe bis jetzt die vielfältigen Einzelheiten der Sache beachtet
Bin bis jetzt unbekümmert vorgegangen	3 2 1 0 1 2 3 0-0-0-0-0-0-0-0-0-0	Bin bis jetzt sorgsam vorgegangen
Habe bis jetzt Dinge miteinander verknüpft	3 2 1 0 1 2 3 0-0-0-0-0-0-0-0-0-0	Habe bis jetzt Dinge voneinander getrennt
Habe bis jetzt Einfälle aufsteigen lassen	3 2 1 0 1 2 3 0-0-0-0-0-0-0-0-0-0	Habe bis jetzt Einfälle unterdrückt

Auf den folgenden Seiten findest Du Items, die Deine jetzigen Gefühle betreffen. Es geht also nicht darum, wie Du Dich in den letzten Monaten im Allgemeinen gefühlt hast, sondern darum, was Du genau jetzt empfindest.

Auf jeder Seite sind die Items von "1" beginnend numeriert. Laß Dich davon nicht stören! Das hat drucktechnische Gründe.

Achte dafür jedoch bitte auf die Antwortkategorien. Sie sind von Seite zu Seite verändert!

Mache Deine Kreuze je nach Skala bitte so:

X---O---O---O---O

oder so:

.....| |X| | |

oder so:

1---2---X---4---5---6---7---8---9

Beantworte bitte jedes Item so, wie es Deiner augenblicklichen Stimmung entspricht. Laß bitte kein Item aus.

Gehe zügig vor! Überlege nicht stundenlang!

Ich erlebe in diesem Moment:

keine
Angst

1-----2-----3-----4-----5-----6-----7-----8-----9

die stärkste
Angst, die ich
je hatte

überhaupt
keine wut

1-----2-----3-----4-----5-----6-----7-----8-----9

so starke Wut,
wie ich sie
bei mir noch
nie bemerkte

kein ein-
ziges 1
depressives
Gefühl

1-2-3-4-5-6-7-8-9

den schlimmsten
depressiven
Zustand, unter
dem ich je litt

kein
Lampen-
fieber

1-----2-----3-----4-----5-----6-----7-----8-----9

so starkes
Lampenfieber,
daß ich mich
an kein stär-
keres erinnern
kann

Ich habe
alles unter
unter 1
Kontrolle

1-2-3-4-5-6-7-8-9

Ich bin völlig hilflos

überhaupt
kein Arger

1-2-3-4-5-6-7-8-9

so starken
Arger, wie ich
ihn noch nie
empfand

	stimmt genau	ganz im Gegenteil
1. Ich bin völlig hilflos.	0-----0-----0-----0-----0	
2. Ich glaube nicht, daß ich es je schaffe.	0-----0-----0-----0-----0	
3. Ich bin unfähig.	0-----0-----0-----0-----0	
4. Ich bin traurig, weil ich nichts machen kann.	0-----0-----0-----0-----0	
5. Ich fühle mich überfordert.	0-----0-----0-----0-----0	
6. Egal, was ich tue, es wird sich nichts ändern.	0-----0-----0-----0-----0	
7. Ich werde alles falsch machen, was von mir erwartet wird.	0-----0-----0-----0-----0	
8. Ich fühle mich ganz mies.	0-----0-----0-----0-----0	
9. Ich kann keinen klaren Gedanken fassen.	0-----0-----0-----0-----0	
10. Ich komme nicht mehr zurecht.	0-----0-----0-----0-----0	
11. Ich bin unglücklich.	0-----0-----0-----0-----0	
12. Ich bin völlig abhängig.	0-----0-----0-----0-----0	
13. Ich bin ausgeliefert.	0-----0-----0-----0-----0	
14. Ich habe keinen Mut.	0-----0-----0-----0-----0	
15. Andere würden im Moment besser zurechtkommen.	0-----0-----0-----0-----0	
16. Meine Situation ist völlig ver- fahren.	0-----0-----0-----0-----0	
17. Meine Hilflosigkeit ist uner- träglich.	0-----0-----0-----0-----0	
18. Ich fühle, daß die anderen stär- ker sind.	0-----0-----0-----0-----0	
19. Selbst wenn ich mich anstrenge, kommt doch nichts heraus.	0-----0-----0-----0-----0	
20. Ich weiß nicht, was ich noch tun soll.	0-----0-----0-----0-----0	
21. Ich bin einsam und allein.	0-----0-----0-----0-----0	
22. Eigentlich macht mir im Moment nichts mehr Freude.	0-----0-----0-----0-----0	

	falsch	stimmt überwiegend nicht	stimmt teilweise	stimmt überwiegend	stimmt genau
1. Ich sehe hoffnungsvoll und enthusiastisch in die Zukunft					
2. Ich könnte ebensogut aufgeben, weil ich meine Situation nicht verbessern kann					
3. Wenn etwas schlecht für mich läuft, hilft mir das Wissen, daß es auch mal besser wird					
4. Ich kann mir nicht vorstellen, wie ich in 10 Jahren lebe					
5. Ich habe genug Zeit, um die Dinge zu tun, die mir am meisten Freude machen					
6. In den Bereichen, die für mich wichtig sind, werde ich Erfolg haben					
7. Für meine Zukunft sehe ich schwarz					
8. Ich glaube, daß es mir im Leben besser gehen wird als vielen anderen					
9. Ich schaffe zur Zeit kaum etwas, und ich sehe keinen Grund, warum sich das zukünftig ändern sollte					
10. Mein bisheriges Leben hat mich auf die Zukunft gut vorbereitet					
11. Alles, was auf mich zukommt, ist unangenehm. Ich erwarte nichts Angenehmes					
12. Was ich mir wirklich wünsche, werde ich vermutlich nicht erreichen					
13. Ich erwarte, in der Zukunft glücklicher als jetzt zu sein					
14. Die Dinge werden sich nicht so entwickeln, wie ich das gern hätte					
15. Ich sehe sehr zuversichtlich in die Zukunft					
16. weil es nie so läuft, wie ich das gern hätte, ist es sinnlos, irgendetwas zu versuchen					
17. Vermutlich werde ich zukünftig keine Erfüllung finden					
18. Meine Zukunft erscheint mir vage und unsicher					
19. Ich sehe eher gute als schlechte Zeiten auf mich zukommen					
20. Es ist zwecklos, etwas für mich Wichtiges anzustreben, weil ich es vermutlich nicht erreiche					

	stimmt genau	ganz im Ge- genteil
1. Es ist mir verleidet, oder: ich bin traurig.	0---0---0---0---0	
2. Ich glaube, daß ich nichts zu erhoffen habe.	0---0---0---0---0	
3. Ich habe nicht das Gefühl, mehr versagt zu haben als andere Leute.	0---0---0---0---0	
4. Ich habe an nichts mehr dieselbe Freude wie früher.	0---0---0---0---0	
5. Ich empfinde mich sehr oft als schlecht oder wertlos.	0---0---0---0---0	
6. Ich glaube, daß mir etwas Schlimmes passieren könnte.	0---0---0---0---0	
7. Ich bin nicht enttäuscht von mir.	0---0---0---0---0	
8. Ich achte sehr gut auf meine eigenen Irrtümer oder Fehler.	0---0---0---0---0	
9. Ich mache mir Vorwürfe für alles, was nicht klappt.	0---0---0---0---0	
10. Ich denke manchmal daran, mir etwas anzutun aber ich werde es nicht tun.	0---0---0---0---0	
11. Ich weine häufiger als gewöhnlich.	0---0---0---0---0	
12. Ich bin rascher verärgert (oder gereizt) als gewöhnlich.	0---0---0---0---0	
13. Alles was mich gewöhnlich geärgert hat, berührt mich nicht mehr.	0---0---0---0---0	
14. Ich interessiere mich weniger als gewöhnlich für andere Leute.	0---0---0---0---0	
15. Ich bin jetzt weniger selbstsicher und versuche, Entscheidungen aus dem Wege zu gehen.	0---0---0---0---0	
16. Ich befürchte, daß ich gealtert oder unvorteilhaft aussehe.	0---0---0---0---0	
17. Es braucht eine zusätzliche Anstrengung, an etwas heranzugehen.	0---0---0---0---0	
18. Ich arbeite nicht mehr so gut wie gewohnt.	0---0---0---0---0	
19. Mein Appetit ist nicht so gut wie gewöhnlich.	0---0---0---0---0	
20. Ich interessiere mich viel weniger fürs Sexuelle	0---0---0---0---0	
21. Ich bin sehr mit Unwohlsein und Schmerzen beschäftigt, zum Beispiel mit Bauchweh oder Verstopfung oder mit anderen unangenehmen körperlichen Empfindungen.	0---0---0---0---0	

	Überhaupt nicht	Ein wenig	Ziemlich	In starkem Ausmaß	In sehr starkem Ausmaß
1. Ich fühle mich ruhig					
2. Ich fühle mich sicher					
3. Ich bin gespannt					
4. Ich bedauere etwas					
5. Ich fühle mich wohl					
6. Ich bin aufgeregt					
7. Ich mache mir zur Zeit Sorgen, daß etwas schief gehen könnte					
8. Ich fühle mich ausgeruht					
9. Ich bin beunruhigt					
10. Ich fühle mich behaglich					
11. Ich habe Selbstvertrauen					
12. Ich fühle mich nervös					
13. Ich bin zappelig					
14. Meine Nerven sind bis zum Zerreißen angespannt					
15. Ich bin entspannt					
16. Ich fühle mich zufrieden					
17. Ich mache mir Sorgen					
18. Ich fühle mich überreizt und durcheinander					
19. Ich fühle mich froh					
20. Ich bin vergnügt					

Du wirst auf den folgenden Seiten eine Reihe von Fragen nach verschiedenen körperlichen Vorgängen und Beschwerden finden. Antworte bitte, wie Du Dich während der letzten Zeit gefühlt hast.

Jeweils wird nach der Häufigkeit der Beschwerden gefragt. Die Antwortmöglichkeiten, die zur Vereinfachung bereits vorgedruckt sind, lauten:

fast	etwa 3 mal	etwa 2 mal	etwa 2 mal	nie
tägl.	in d. Woche	im Monat	im Jahr	

Bei diesen Items bitte das Kreuz auf die Linie aufsetzen!
Beispiel:

Herzschmerzen falsch  richtig

wie häufig sind die folgenden Symptome während der letzten Zeit aufgetreten?

[illegible]

	fast tagl.	etwa 3 x Woche	etwa 2 x Monat	etwa 2 x Jahr	nie
31. Nasenjucken, ohne erkältet zu sein					
32. Niesen, ohne erkältet zu sein ..					
33. Halsschmerzen, ohne erkältet zu sein					
34. Schwierigkeiten beim Schlucken..					
35. Gefühl, einen Kloß im Hals zu haben					
36. Räuspern, ohne erkältet zu sein..					
37. Husten, ohne erkältet zu sein...					
38. Kopffucken					
39. Feuchte Hände					
40. Schweißausbrüche					
41. Lichtempfindlichkeit, so daß Sie z.B. von hellem Licht ge- stört werden					
42. Farbempfindlichkeit, so daß Sie z.B. von krassen Farben oder Farbgegensätzen gestört wer- den					
43. Geräuschempfindlichkeit, so daß Sie z.B. von lauten Geräuschen und Tönen gestört werden					
44. Geruchsempfindlichkeit, so daß z.B. bestimmte Gerüche Sie stören					
45. Geschmacksempfindlichkeit, so daß Sie z.B. Ihr Essen nur schwach würzen					
46. Empfindlichkeit gegenüber unbe- quemer, enger Kleidung					
47. Druck hinter den Augen					
48. Nackenschmerzen					
49. Schulterschmerzen					
50. Kreuzschmerzen					
51. Schmerzen in den Armen					
52. Mißempfindungen an den Armen oder Händen (Kribbeln, Prik- keln, Taubheit)					
53. Mißempfindungen an den Beinen oder Füßen (Kribbeln, Prik- keln, Taubheit)					
54. Schmerzen in den Beinen					
55. Ganzer Körper verkrampft					

	fast tagl.	etwa 3 x Woche	etwa 2 x Monat	etwa 2 x Jahr	nie
56. Verkrampfen während des Arbeitens					
57. Verkrampfen beim Schreiben.....					
58. Ungeschickt hantieren					
59. Zitterige Hände häufig beim Anzünden einer Zigarette oder beim Halten einer Tasse					
60. Ruhelose Hände					
61. Ruhelose Füße					
62. Unwillkürliches Zucken um das Auge					
63. Unwillkürliches Zucken um den Mund					
64. Unwillkürliches Schnalzen mit der Zunge					
65. Unwillkürliches Zucken der Schultern					
66. Empfindlichkeit gegen Kälte.....					
67. Kalte Hände selbst bei warmer Witterung					
68. Empfindliche Haut					
69. Trockene Haut					
70. Eine Allergie, d.h. eine spezielle Überempfindlichkeitsreaktion (z.B. Heuschnupfen, Ekzem)					
71. Schmerzempfindlichkeit					

Bei den folgenden Items handelt es sich um Statements, die auf die Situation der Prüfung zugeschnitten sind. Wenn Du über die angeschnittene Möglichkeit noch nicht recht nachgedacht hast, dann kreuze bitte "noch nie darüber nachgedacht" an. Solltest Du die Möglichkeit bereits überlegt haben, so kannst Du das Statement entsprechend Deiner augenblicklichen Prüfungsstrategie für Dich ablehnen ("eher falsch" und "falsch") oder annehmen ("eher richtig" und "richtig"). Solltest Du Dich nicht entscheiden können, dann kreuze bitte "starkes Problem für mich, weiß ich aber nicht" an. Arbeite genau, aber zügig!

	richtig	eher richtig	starkes Problem für mich, weiß ich aber nicht	eher falsch	falsch	noch nie darüber nachgedacht
1. Wenn ich anderer Meinung bin, dann werde ich auch in einer Prüfung einem Professor widersprechen.						
	0-----	0-----	0-----	0-----	0-----	/ 0
2. Ich werde betont selbstbewußt auftreten.						
	0-----	0-----	0-----	0-----	0-----	/ 0
3. Ich werde mich optimistisch zeigen.						
	0-----	0-----	0-----	0-----	0-----	/ 0
4. Ich werde möglichst wenig sagen.						
	0-----	0-----	0-----	0-----	0-----	/ 0
5. Ich werde mich mutig zeigen.						
	0-----	0-----	0-----	0-----	0-----	/ 0
6. Ich werde mich vorsichtig verhalten.						
	0-----	0-----	0-----	0-----	0-----	/ 0
7. Ich konstruiere The- sen.						
	0-----	0-----	0-----	0-----	0-----	/ 0
8. Am Anfang bin ich sicherlich aufgeregt, aber das wird sich dann legen.						
	0-----	0-----	0-----	0-----	0-----	/ 0
9. Wahrscheinlich wird es vorkommen, daß mir eine Antwort vor und nach der Prüfung einfällt, ich sie während der Prüfung aber nicht weiß.						
	0-----	0-----	0-----	0-----	0-----	/ 0
10. Wahrscheinlich werde ich mich gegen Ende der Prüfung kaum noch konzentrieren können.						
	0-----	0-----	0-----	0-----	0-----	/ 0
11. Ich muß alle Fragen sehr genau und gewissenhaft beantworten.						
	0-----	0-----	0-----	0-----	0-----	/ 0

				starkes Problem für mich, weiß ich aber nicht	eher falsch	falsch	noch nie darüber nachgedacht
	richtig	richtig					
12. Vor jeder Antwort überlege ich genau die Formulierung	0-----	0-----	0-----	0-----	0-----	0	/ 0
13. Es ist in jedem Fall besser, wenn ich erst einmal tief durchatme, statt spontan mit der Antwort zu beginnen.	0-----	0-----	0-----	0-----	0-----	0	/ 0
14. Das Beste ist, wenn ich kurz und bündig auf die Fragen des Prüfers antworte.	0-----	0-----	0-----	0-----	0-----	0	/ 0
15. Wenn ich in einer Prüfung viel rede, erzähle ich auch viel Falsches.	0-----	0-----	0-----	0-----	0-----	0	/ 0
16. Ich kann mir vorstellen, daß ein Prüfer versucht, mich fertigzumachen.	0-----	0-----	0-----	0-----	0-----	0	/ 0
17. Wenn ich auf eine Frage die Antwort nicht weiß, dann heißt das, daß ich bei der Vorbereitung nicht aufgepaßt habe.	0-----	0-----	0-----	0-----	0-----	0	/ 0
18. Wenn ein Prüfer mir eine Frage stellt, deren Antwort ich überhaupt nicht kenne, dann prüft er mich auf "sehr gut".	0-----	0-----	0-----	0-----	0-----	0	/ 0
19. In einem gewissen Rahmen kann ich auf den Prüfungsverlauf Einfluß nehmen.	0-----	0-----	0-----	0-----	0-----	0	/ 0
20. Es hat auf die Prüfung keinerlei Einfluß, wenn ich zugebe, daß ich nicht alles weiß.	0-----	0-----	0-----	0-----	0-----	0	/ 0
21. Die Zensur interessiert mich überhaupt nicht, weil ich lediglich bestehen möchte.	0-----	0-----	0-----	0-----	0-----	0	/ 0
22. Ich gehe davon aus, daß mir in der Prüfung auch auf schwere Fragen eine Antwort einfallen wird.	0-----	0-----	0-----	0-----	0-----	0	/ 0
23. Im großen und ganzen ist mir bekannt, welcher Stoff geprüft werden wird.	0-----	0-----	0-----	0-----	0-----	0	/ 0

				starkes Problem für mich, weiß ich aber nicht	eher falsch	falsch	noch nie darüber nachgedacht
	richtig	eher richtig	richtig	nicht	eher falsch	falsch	nachgedacht
24. Prüfer stellen nur scheinbar leichte Fragen, in Wirklichkeit wollen sie mehr wissen.	0-----	0-----	0-----	0-----	0-----	0	/ 0
25. Detailkenntnisse sind mindestens genauso wichtig wie der Überblick.	0-----	0-----	0-----	0-----	0-----	0	/ 0
26. Es kommt darauf an, Zusammenhänge zu sehen und zu beschreiben.	0-----	0-----	0-----	0-----	0-----	0	/ 0
27. Wichtig ist, auf den Kern der Sache zu kommen. Die Einzelheiten drumherum kann ich vernachlässigen.	0-----	0-----	0-----	0-----	0-----	0	/ 0
28. Es kommt darauf an, daß ich die Zusammenhänge und Unterschiede anschaulich darstellen.	0-----	0-----	0-----	0-----	0-----	0	/ 0
29. Es ist für die Zensur unerheblich, wenn ich ein paar Details nicht weiß.	0-----	0-----	0-----	0-----	0-----	0	/ 0
30. In den Prüfungen wird niemals die Fähigkeit, sondern der Fleiß geprüft.	0-----	0-----	0-----	0-----	0-----	0	/ 0
31. Es ist besser, wenn ich nicht allzu kritisch auftrete.	0-----	0-----	0-----	0-----	0-----	0	/ 0
32. Wenn ich an einer Theorie etwas aussetzen habe, werde ich dies sehr deutlich zum Ausdruck bringen.	0-----	0-----	0-----	0-----	0-----	0	/ 0
33. Es ist wichtig, daß ich spontan antworte.	0-----	0-----	0-----	0-----	0-----	0	/ 0
34. Wenn ich bei mehr als 3 Fragen überlegen muß, macht das einen schlechten Eindruck.	0-----	0-----	0-----	0-----	0-----	0	/ 0

Bestimmt hast Du Dir schon Gedanken gemacht, wie Du Dich in den mündlichen Prüfungen verhalten wirst. Darum geht es bei den folgenden Items.

Dazu sind 15 Aussagenalternativen formuliert. Zwischen den Alternativen befindet sich eine Skala. Mache Dein Kreuz in die linke Hälfte, wenn die linke Aussagenalternative zutrifft. Wähle die rechte Hälfte, wenn die rechte Alternative zutrifft. Kreuze bitte die "3" an, wenn die Aussage genau zutrifft. Trifft eine der Alternativen eher zu als die andere, so kreuze bitte die "1" oder "2" an. Die "unentschiedene" Mittelposition - die "0" - wähle bitte nur dann, wenn für Dich eine Entscheidung völlig unmöglich ist. Wie wirst Du Dich in den mündlichen Prüfungen verhalten?

werde phantasievoll denken	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde nüchtern denken
werde besonnen vorgehen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde zügig vorgehen
Werde den Gegenstand ganzheitlich auffassen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde den Gegenstand zergliedern
Werde die Zusammenhänge auflösen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde Zusammenhänge herstellen
Werde leichtfüßig vorgehen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde hartnäckig vorgehen
Werde die Sache in den Griff nehmen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde die Sache auf mich wirken lassen
Werde das Verhältnis der Einzelheiten zu einem übergeordneten Ganzen wichtig nehmen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	werde das Verhältnis der Einzelheiten untereinander betonen
Werde rasch vorgehen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	werde gründlich vorgehen
Werde systematisch vorgehen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde ohne strenge Systematik vorgehen
Werde variabel vorgehen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde stetig vorgehen
Werde mich anstrengen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde gelassen bleiben
Werde den Kern der Sache beachten	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde die vielfältigen Einzelheiten der Sache beachten
Werde unbekümmert vorgehen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	werde sorgsam vorgehen
Werde Dinge miteinander verknüpfen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	werde Dinge voneinander trennen
werde Einfälle aufsteigen lassen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde Einfälle unterdrücken

PS.: Für Deine Mitarbeit danke ich Dir, auch im Namen von noch unbekannten Prüflingen.

Für die bevorstehenden Prüfungen wünsche ich Dir alles Gute!

Bochum, den 14.1.1976

Lodhar Breidkopf

7.2 Fragebogen AN

Bearbeitung nach der Prüfung vor Bekanntgabe der Zensur

Der hier wiedergegebene Fragebogen gliedert sich wie folgt:

	Seite
1 Seite Deckblatt (wurde nach Bearbeitung vernichtet)	500
1 Seite Fragen nach der erwarteten Zensur, der Reaktion auf eine Nachprüfung, der momentanen Erschöpfung, der momentanen Konzentrationsfähigkeit, dem Prüfungsverlauf und dem Prozentsatz des geprüften Stoffes	501
1 Seite Analogskalen	502

AN

(Projekt "Prüfungsstress")

Lothar Breitkopf

Dieses Deckblatt wird nach Umcodierung der Matrikelnummer
abgerissen und vernichtet.

Trage hier bitte Deine Matrikelnummer (oder den im 1. Frage-
bogen verwendeten Code) ein:

CODE ▶ ▶ ▶

--	--	--	--	--	--

► Bei wem hast Du die Prüfung eben abgelegt?

- ☐ Heckhausen
- ☐ Hörmann
- ☐ Graefe
- ☐ Schönbach
- ☐ Schulte

► Welche Zensur wirst Du bekommen (diese Frage in jedem Fall beantworten; ggf. raten!) ?

1-----	2-----	3-----	4-----	5-----
sehr	gut	befrie-	ausrei-	durchge-
gut		digend	chend	fallen

► Die wievielte von allen mündlichen Prüfungen, die Du insgesamt ablegen mußt, war die soeben beendete?

1.-----2.-----3.-----4.-----5.-----6.-----7.

► Gesetzt den Fall, der Prüfer würde Dir gleich die Chance geben, Deine Zensur durch eine Nachprüfung (5 Min.) jetzt sofort zu verbessern. Wie würdest Du reagieren ?

spontan	0----	0----	0----	0----	0----	0----	0----	0----	0----	spontan
zustimmen										ablehnen

► Wie erschöpft fühlst Du Dich in diesem Moment?

fühle										so erschöpft
mich	1----	2----	3----	4----	5----	6----	7----	8----	9----	fühlte ich
voll fit										mich noch nie

► Wie schätzt Du Deine momentane Konzentrationsfähigkeit ein?

So hoch										so niedrig
wie noch	0----	0----	0----	0----	0----	0----	0----	0----	0----	wie noch
nie										nie

► War der Verlauf der Prüfung genau so, wie Du Dir das vorgestellt hattest?

nein, viel										nein, viel
schlechter	0----	0----	0----	0----	0----	0----	0----	0----	0----	besser

► Mit welchem Prozentsatz des geprüften Stoffes hast Du Dich bereits während des studiums (=frühere Semester) beschäftigt?

0-----	10-----	25-----	50-----	75-----	90-----	100-----
0%	10%	25%	50%	75%	90%	100%

Was erlebst Du genau in diesem Moment (also: wenn Du diese Zeilen liest) ?

keine Angst	1-----2-----3-----4-----5-----6-----7-----8-----9	die stärkste Angst, die ich je hatte
-------------	---	--------------------------------------

überhaupt keine Wut	1-----2-----3-----4-----5-----6-----7-----8-----9	so starke Wut, wie ich sie bei mir noch nie bemerkte
---------------------	---	--

kein einziges depressives Gefühl	1-----2-----3-----4-----5-----6-----7-----8-----9	den schlimmsten depressiven Zustand, unter dem ich je litt
----------------------------------	---	--

kein Lampenfieber	1-----2-----3-----4-----5-----6-----7-----8-----9	so starkes Lampenfieber, daß ich mich an kein stärkeres erinnern kann
-------------------	---	---

keine Freude	1-----2-----3-----4-----5-----6-----7-----8-----9	extrem stark Freude
--------------	---	---------------------

Ich habe alles unter Kontrolle	1-----2-----3-----4-----5-----6-----7-----8-----9	Ich bin völlig hilflos
--------------------------------	---	------------------------

überhaupt kein Arger	1-----2-----3-----4-----5-----6-----7-----8-----9	so starken Arger, wie ich ihn noch nie empfand
----------------------	---	--

7.3 Fragebogen F2

Bearbeitung nach Bekanntgabe der Zensur

Der hier wiedergegebene Fragebogen gliedert sich wie folgt:

	Seite
1 Seite Deckblatt (wurde nach Bearbeitung vernichtet)	504
1 Seite Zensur, die emotionale Reaktion auf die Zensur und ihre Kausalattributionierung	505
1 Seite Analogskalen ' <u>vor</u> der Prüfung'	506
Emotionstests 'während der Prüfung'	
1 Seite H1S	508
1 Seite BDI	509
1 Seite STAI	510
1 Seite Analogskalen 'während der Prüfung'	511
1 Seite Beschwerdeliste	512
Strategierealisation	
1 Seite Instruktion	513
1 Seite TVD-Habe	514

F 2

(Projekt "Prüfungsstress")

Lothar Breitkopf

Dieses Deckblatt wird nach Umcodierung der Matrikelnummer
abgerissen und vernichtet.

Trage hier bitte Deine Matrikelnummer (oder den im 1. Frage-
bogen verwendeten Identifikationscode) ein:

CODE ►►►

--	--	--	--	--	--

► Welche Zensur hast Du bekommen?

► Welche Gefühle löst die Zensur in Dir aus ?
(Bitte alles Zutreffende ankreuzen!)

- ☐ Erleichterung
- ☐ Freude
- ☐ Ärger
- ☐ Enttäuschung
- ☐ Wut
- ☐ Angst
- ☐ Depressive Gefühle
- ☐ Zufriedenheit
- ☐ Zuversicht
- ☐ Sicherheit
- ☐ Verwirrung
- ☐ Hilflosigkeit

► worauf führst Du Dein Abschneiden zurück? In welchem Ausmaß siehst Du die Gründe dafür in den folgenden Ursachen?

	bestimmt	vermut-	vielleicht	ziemlich	völlig
	nicht	nicht	doch	sicher	sicher
1. Ich habe mich angestrengt.	0-----	0-----	0-----	0-----	0-----
2. Ich hatte Glück.	0-----	0-----	0-----	0-----	0-----
3. Es liegt an meinen Fähigkeiten.	0-----	0-----	0-----	0-----	0-----
4. Der Prüfer prüft gut.	0-----	0-----	0-----	0-----	0-----
5. Ich war faul.	0-----	0-----	0-----	0-----	0-----
6. Ich hatte Pech.	0-----	0-----	0-----	0-----	0-----
7. Das liegt an meiner Unfähigkeit.	0-----	0-----	0-----	0-----	0-----
8. Der Prüfer prüft schlecht.	0-----	0-----	0-----	0-----	0-----
9. Ich war fleißig.	0-----	0-----	0-----	0-----	0-----
10. Das war Zufall.	0-----	0-----	0-----	0-----	0-----
11. Ich bin intelligent.	0-----	0-----	0-----	0-----	0-----
12. Es wird nicht viel verlangt.	0-----	0-----	0-----	0-----	0-----

- Mit den folgenden Skalen sollen Deine Gefühle unmittelbar vor
 ► der Prüfung erfaßt werden. Vergewähre Dir dazu bitte, wie
 ► Du gewartet, was Du empfunden hast, bevor Du das Prüfungszimmer
 ► betreten hast. Wie hättest Du in jenem Moment geantwortet?

keine Angst 1----2----3----4----5----6----7----8----9 die stärkste Angst, die ich je hatte

überhaupt keine Wut 1----2----3----4----5----6----7----8----9 so starke Wut, wie ich sie bei mir noch nie bemerkte

kein einziges depressives Gefühl 1----2----3----4----5----6----7----8----9 den schlimmsten depressiven Zustand, unter dem ich je litt

kein Lampenfieber 1----2----3----4----5----6----7----8----9 so starkes Lampenfieber, daß ich mich an kein stärkeres erinnern kann

keine Freude 1----2----3----4----5----6----7----8----9 extrem starke Freude

Ich habe alles unter Kontrolle 1----2----3----4----5----6----7----8----9 Ich bin völlig hilflos

überhaupt kein Ärger 1----2----3----4----5----6----7----8----9 so starken Ärger, wie ich ihn noch nie empfand

Auf den folgenden Seiten geht es um Deine Gefühle und Empfindungen während der Prüfung. wie hast Du Dich die meiste Zeit gefühlt?

Bearbeite bitte jedes Item! Gehe genau, aber zügig vor! Aus drucktechnischen Gründen sind die Items auf jeder Seite neu nummeriert. Laß Dich dadurch bitte nicht verwirren.

Wie hättest Du während der Prüfung geantwortet ?

	stimmt genau	ganz im Gegenteil
1. Ich bin völlig hilflos.	0-----0-----0-----0-----0	
2. Ich glaube nicht, daß ich es je schaffe.	0-----0-----0-----0-----0	
3. Ich bin unfähig.	0-----0-----0-----0-----0	
4. Ich bin traurig, weil ich nichts machen kann.	0-----0-----0-----0-----0	
5. Ich fühle mich überfordert.	0-----0-----0-----0-----0	
6. Egal, was ich tue, es wird sich nichts ändern.	0-----0-----0-----0-----0	
7. Ich werde alles falsch machen, was von mir erwartet wird.	0-----0-----0-----0-----0	
8. Ich fühle mich ganz mies.	0-----0-----0-----0-----0	
9. Ich kann keinen klaren Gedanken fassen.	0-----0-----0-----0-----0	
10. Ich komme nicht mehr zurecht.	0-----0-----0-----0-----0	
11. Ich bin unglücklich.	0-----0-----0-----0-----0	
12. Ich bin völlig abhängig.	0-----0-----0-----0-----0	
13. Ich bin ausgeliefert.	0-----0-----0-----0-----0	
14. Ich habe keinen Mut.	0-----0-----0-----0-----0	
15. Andere würden im Moment besser zurechtkommen.	0-----0-----0-----0-----0	
16. Meine Situation ist völlig ver- fahren.	0-----0-----0-----0-----0	
17. Meine Hilflosigkeit ist uner- träglich.	0-----0-----0-----0-----0	
18. Ich fühle, daß die anderen stür- ker sind.	0-----0-----0-----0-----0	
19. Selbst wenn ich mich anstrenge, kommt doch nichts heraus.	0-----0-----0-----0-----0	
20. Ich weiß nicht, was ich noch tun soll.	0-----0-----0-----0-----0	
21. Ich bin einsam und allein.	0-----0-----0-----0-----0	
22. Eigentlich macht mir im Moment nichts mehr Freude.	0-----0-----0-----0-----0	

Wie hättest Du während der Prüfung geantwortet?

	stimmt genau	ganz im Ge- gentell
1. Es ist mir verleidet, oder: ich bin traurig.	0-----0-----0-----0-----0	
2. Ich glaube, daß ich nichts zu erhoffen habe.	0-----0-----0-----0-----0	
3. Ich habe nicht das Gefühl, mehr versagt zu haben als andere Leute.	0-----0-----0-----0-----0	
4. Ich habe an nichts mehr dieselbe Freude wie früher.	0-----0-----0-----0-----0	
5. Ich empfinde mich sehr oft als schlecht oder wertlos.	0-----0-----0-----0-----0	
6. Ich glaube, daß mir etwas Schlimmes passieren könnte.	0-----0-----0-----0-----0	
7. Ich bin nicht enttäuscht von mir.	0-----0-----0-----0-----0	
8. Ich achte sehr gut auf meine eigenen Irrtümer oder Fehler.	0-----0-----0-----0-----0	
9. Ich mache mir Vorwürfe für alles, was nicht klappt.	0-----0-----0-----0-----0	
10. Ich denke manchmal daran, mir etwas anzutun aber ich werde es nicht tun.	0-----0-----0-----0-----0	
11. Ich weine häufiger als gewöhnlich.	0-----0-----0-----0-----0	
12. Ich bin rascher verärgert (oder gereizt) als gewöhnlich.	0-----0-----0-----0-----0	
13. Alles was mich gewöhnlich geärgert hat, berührt mich nicht mehr.	0-----0-----0-----0-----0	
14. Ich interessiere mich weniger als gewöhnlich für andere Leute.	0-----0-----0-----0-----0	
15. Ich bin jetzt weniger selbstsicher und versuche, Entscheidungen aus dem Wege zu gehen.	0-----0-----0-----0-----0	
16. Ich befürchte, daß ich gealtert oder unvorteilhaft aussehe.	0-----0-----0-----0-----0	
17. Es braucht eine zusätzliche Anstrengung, an etwas heranzugehen.	0-----0-----0-----0-----0	
18. Ich arbeite nicht mehr so gut wie gewohnt.	0-----0-----0-----0-----0	
19. Mein Appetit ist nicht so gut wie gewöhnlich.	0-----0-----0-----0-----0	
20. Ich interessiere mich viel weniger fürs Sexuelle	0-----0-----0-----0-----0	
21. Ich bin sehr mit Unwohlsein und Schmerzen beschäftigt, zum Beispiel mit Bauchweh oder Verstopfung oder mit anderen unangenehmen körperlichen Empfindungen.	0-----0-----0-----0-----0	

Wie hättest Du während der Prüfung geantwortet?

	überhaupt nicht	Bin wenig	Ziemlich	In starkem Ausmaß	In sehr starkem Ausmaß
1. Ich fühle mich ruhig					
2. Ich fühle mich sicher					
3. Ich bin gespannt					
4. Ich bedauere etwas					
5. Ich fühle mich wohl					
6. Ich bin aufgeregt					
7. Ich mache mir zur Zeit Sorgen, daß etwas schief gehen könnte					
8. Ich fühle mich ausgeruht					
9. Ich bin beunruhigt					
10. Ich fühle mich behaglich					
11. Ich habe Selbstvertrauen					
12. Ich fühle mich nervös					
13. Ich bin zappelig					
14. Meine Nerven sind bis zum Zerreißen angespannt					
15. Ich bin entspannt					
16. Ich fühle mich zufrieden					
17. Ich mache mir Sorgen					
18. Ich fühle mich überreizt und durcheinander					
19. Ich fühle mich froh					
20. Ich bin vergnügt					

Ich empfand während der Prüfung

keine		die stärkste
Angst	1----2----3----4----5----6----7----8----9	Angst, die ich je hatte

überhaupt		so starke Wut,
keine Wut	1----2----3----4----5----6----7----8----9	wie ich sie bei mir noch nie bemerkte

kein ein-		den schlimmsten
ziges	1----2----3----4----5----6----7----8----9	depressiven Zustand, unter dem ich je litt
depressives		
Gefühl		

kein		so starkes
Lampen-	1----2----3----4----5----6----7----8----9	Lampenfieber,
fieber		daß ich mich an kein stärkeres erinnern kann

keine		extrem starke
Freude	1----2----3----4----5----6----7----8----9	Freude

Ich habe		
alles unter	1----2----3----4----5----6----7----8----9	Ich bin völlig hilflos
Kontrolle		

überhaupt	1----2----3----4----5----6----7----8----9	so starken
kein Ärger		Ärger, wie ich ihn noch nie empfand

► Wie häufig traten die folgenden Symptome bei Lir während der Prüfung auf?

	fast ständig	häu- fig	1 - 2 mal	nie
- Müdigkeit und Zerschlagenheit	o-----o-----o-----o			
- Überraschendes Ermüden	o-----o-----o-----o			
- Schwindelgefühle	o-----o-----o-----o			
- Benommenheit	o-----o-----o-----o			
- Kopfschmerzen	o-----o-----o-----o			
- Gefühl, gleich weinen zu müssen	o-----o-----o-----o			
- Stottern	o-----o-----o-----o			
- Erröten	o-----o-----o-----o			
- Atembeschwerden	o-----o-----o-----o			
- Anfallsartiges Herzklopfen	o-----o-----o-----o			
- Gefühl, auf die Toilette zu müssen	o-----o-----o-----o			
- Zittern oder "weiche knie"	o-----o-----o-----o			
- Feuchte Hände	o-----o-----o-----o			
- Hitzewellen	o-----o-----o-----o			
- Schwitzen	o-----o-----o-----o			
- Verkrampfungen im Bauchraum	o-----o-----o-----o			
- Verkrampfungen im Nacken	o-----o-----o-----o			
- Übelkeit	o-----o-----o-----o			

Bestimmt kannst Du Dich noch erinnern, wie Du Dich in der Prüfung verhalten hast. Darum geht es auf der folgenden Seite.

Du findest 15 Aussagenpaare formuliert. Zwischen jedem Paar befindet sich eine Skala. Mache Dein Kreuz in die linke Hälfte, wenn die linke Aussagenalternative zutrifft. Wähle die rechte Hälfte, wenn die rechte Alternative zutrifft. Kreuze bitte die "3" an, wenn die Aussage genau zutrifft. Trifft eine der Alternativen eher zu als die andere, so kreuze bitte die "1" oder "2" an. Die "unentschiedene" Mittelposition - die "0" - wähle bitte nur dann, wenn für Dich eine Entscheidung völlig unmöglich ist. Wie hast Du Dich in dieser Prüfung verhalten?

Habe jetzt phantasievoll gedacht	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Habe jetzt nüchtern gedacht
Bin jetzt besonnen vorgegangen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Bin jetzt zügelig vorgegangen
Habe jetzt den Gegen- stand ganzheitlich aufgefaßt	3 2 1 0 1 2 2 0--0--0--0--0--0--0	Habe jetzt den Gegen- stand zergliedert
Habe diesmal die Zu- sammenhänge aufgelöst	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Habe diesmal Zusammen- hänge hergestellt
Bin jetzt leichttrübig vorgegangen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Bin jetzt hartnäckig vorgegangen
Habe jetzt die Sache in den Griff genommen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Habe jetzt die Sache auf mich wirken lassen
Habe jetzt das Verhältnis der Einzelheiten zu einer übergeordneten Ganzen wichtig genommen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Habe jetzt das Ver- hältnis der Einzel- heiten untereinander betont
Bin diesmal rasch vorgegangen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Bin diesmal gründlich vorgegangen
Bin jetzt systematisch vorgegangen	3 2 1 0 1 2 1 0--0--0--0--0--0--0	Bin jetzt ohne strenge Systematik vorgegangen
Bin diesmal variabel vorgegangen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Bin diesmal stetig vorgegangen
Habe mich diesmal angestrengt	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Bin diesmal gelassen geblieben
Habe jetzt den Kern der Sache beachtet	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Habe jetzt die viel- faltigen Einzelheiten der Sache beachtet
Bin diesmal unbekümmert vorgegangen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Bin diesmal sorgsam vorgegangen
Habe jetzt Dinge mittein- ander verknüpft	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Habe jetzt Dinge von- einander trennt
Habe jetzt Einfälle auf- steigen lassen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Habe jetzt Einfälle unterdrückt

7.4 Fragebogen F3

Bearbeitung nach F2

Der hier wiedergegebene Fragebogen gliedert sich wie folgt:

	Seite
1 Seite Deckblatt (wurde nach Bearbeitung vernichtet)	516

1 Seite Erläuterungen	517
--------------------------	-----

Prüfungsstrategien

3 Seiten SS	518-520
----------------	---------

1 Seite TVD-Werde	521
----------------------	-----

F3

(Projekt "Prüfungsstress")

Lothar Breitzkopf

Dieses Deckblatt wird nach Umcodierung der Matrikelnummer
abgerissen und vernichtet.

Trage hier bitte Deine Matrikelnummer (oder den im 1. Fragebogen
verwendeten Identifikationscode) ein:

CODE



--	--	--	--	--	--

Für das Projekt wäre es besser, wenn Du diesen Fragebogen sofort
ausfüllen würdest.

Wenn Du Dich jedoch wirklich nicht in der Lage fühlst, den Frage-
bogen zu bearbeiten, dann nimm ihn mit nach Hause. In diesem Falle
fülle ihn bitte binnen 24 Stunden aus. Wenn Du diese Frist über-
schreitest, werden diese Ergebnisse von Stunde zu Stunde weniger
valide (= ungenauer).

Ich möchte Dich bitten, auch noch diesen Fragebogen jetzt gleich
auszufüllen!

Bei den folgenden Items handelt es sich um Statements, die auf Deine Einschätzung der künftigen Prüfungen bezogen sind.

Wenn Du über die angeschnittene Möglichkeit noch nicht recht nachgedacht hast, dann kreuze bitte ganz rechts "noch nie darüber nachgedacht" an. Solltest Du die Möglichkeit bereits überlegt haben, so kannst Du die Statements entsprechend Deiner momentanen Einschätzung von künftigen Prüfungen ablehnen ("eher falsch" und "falsch") oder annehmen ("eher richtig" und "richtig"). Wenn Du Dich nicht entscheiden kannst, das Statement für Dich aber dennoch wichtig ist, dann kreuze bitte "starkes Problem für mich, weiß ich aber nicht" an.

Bearbeite bitte alle Items! Arbeite genau, aber zügig!

Falls Du diesen Fragebogen mit nach Hause genommen hast:

Wieviele volle Stunden sind seit Deiner letzten Prüfung vergangen?

--	--

Hast Du zwischen der Prüfung und dem jetzigen Augenblick geschlafen?

☐ nein

☐ ja, Mittagsschlaf

☐ ja, Nachtschlaf

☐ ja, ...

(Bitte eintragen!)

Welches Datum haben wir heute (der Tag, an dem Du diese Seite liest)?

--	--

 .

--

 . 1976

was denkst du über Deine künftigen Prüfungen ?



	rich- tig	eher rich- tig	starkes Pro- blem für mich, weiß ich aber nicht	eher falsch	falsch	noch nie darü- ber nach- ge- dacht
1. Wenn ich anderer Meinung bin, dann werde ich auch in einer Prüfung einem Professor widersprechen.	0-----	0-----	0-----	0-----	0-----	/ 0
2. Ich werde betont selbstbewußt auftreten.	0-----	0-----	0-----	0-----	0-----	/ 0
3. Ich werde mich optimistisch zeigen.	0-----	0-----	0-----	0-----	0-----	/ 0
4. Ich werde möglichst wenig sagen.	0-----	0-----	0-----	0-----	0-----	/ 0
5. Ich werde mich mutig zeigen.	0-----	0-----	0-----	0-----	0-----	/ 0
6. Ich werde mich vorsichtig verhalten.	0-----	0-----	0-----	0-----	0-----	/ 0
7. Ich konstruiere The- sen.	0-----	0-----	0-----	0-----	0-----	/ 0
8. Am Anfang bin ich si- cherlich aufgeregt, aber das wird sich dann legen.	0-----	0-----	0-----	0-----	0-----	/ 0
9. Wahrscheinlich wird es vorkommen, daß mir eine Antwort vor und nach der Prüfung einfällt, ich sie während der Prüfung aber nicht weiß.	0-----	0-----	0-----	0-----	0-----	/ 0
10. Wahrscheinlich werde ich mich gegen Ende der Prüfung kaum noch kon- zentrieren können.	0-----	0-----	0-----	0-----	0-----	/ 0
11. Ich muß alle Fragen sehr genau und gewis- senhaft beantworten.	0-----	0-----	0-----	0-----	0-----	/ 0

starkes Problem für

eher mich, weiß

rich- rich- ich aber
tig tig nicht

eher falsch falsch/ge-dacht

- | | tig | tig | nicht | falsch/dacht | |
|---|--------|--------|--------|--------------|-----|
| 2. Vor jeder Antwort überlege ich genau die Formulierung | 0----- | 0----- | 0----- | 0----- | / 0 |
| 3. Es ist in jedem Fall besser, wenn ich erst einmal tief durchatme, statt spontan mit der Antwort zu beginnen. | 0----- | 0----- | 0----- | 0----- | / 0 |
| 4. Das Beste ist, wenn ich kurz und bündig auf die Fragen des Prüfers antworte. | 0----- | 0----- | 0----- | 0----- | / 0 |
| 5. Wenn ich in einer Prüfung viel rede, erzähle ich auch viel Falsches. | 0----- | 0----- | 0----- | 0----- | / 0 |
| 6. Ich kann mir vorstellen, daß ein Prüfer versucht, mich fertigzumachen. | 0----- | 0----- | 0----- | 0----- | / 0 |
| 7. Wenn ich auf eine Frage die Antwort nicht weiß, dann heißt das, daß ich bei der Vorbereitung nicht aufgepaßt habe. | 0----- | 0----- | 0----- | 0----- | / 0 |
| 8. Wenn ein Prüfer mir eine Frage stellt, deren Antwort ich überhaupt nicht kenne, dann prüft er mich auf "sehr gut". | 0----- | 0----- | 0----- | 0----- | / 0 |
| 9. In einem gewissen Rahmen kann ich auf den Prüfungsverlauf Einfluß nehmen. | 0----- | 0----- | 0----- | 0----- | / 0 |
| 10. Es hat auf die Prüfung keinerlei Einfluß, wenn ich zugebe, daß ich nicht alles weiß. | 0----- | 0----- | 0----- | 0----- | / 0 |
| 11. Die Zensur interessiert mich überhaupt nicht, weil ich lediglich bestehen möchte. | 0----- | 0----- | 0----- | 0----- | / 0 |
| 12. Ich gehe davon aus, daß mir in der Prüfung auch auf schwere Fragen eine Antwort einfallen wird. | 0----- | 0----- | 0----- | 0----- | / 0 |
| 13. Im großen und ganzen ist mir bekannt, welcher Stoff geprüft werden wird. | 0----- | 0----- | 0----- | 0----- | / 0 |

				starkes Problem für mich, weiß ich aber nicht			noch nie darüber nachgedacht
24.	Prüfer stellen nur scheinbar leichte Fragen, in Wirklichkeit wollen sie mehr wissen.	richtig	eher richtig	eher falsch	falsch		/ 0
25.	Detaillkenntnisse sind mindestens genauso wichtig wie der Überblick.						/ 0
26.	Es kommt darauf an, Zusammenhänge zu sehen und zu beschreiben.						/ 0
27.	Wichtig ist, auf den Kern der Sache zu kommen. Die Einzelheiten drumherum kann ich vernachlässigen.						/ 0
28.	Es kommt darauf an, daß ich die Zusammenhänge und Unterschiede anschaulich darstelle.						/ 0
29.	Es ist für die Zensur unerheblich, wenn ich ein paar Details nicht weiß.						/ 0
30.	In den Prüfungen wird niemals die Fähigkeit, sondern der Fleiß geprüft.						/ 0
31.	Es ist besser, wenn ich nicht allzu kritisch auftrete.						/ 0
32.	Wenn ich an einer Theorie etwas aussetzen habe, werde ich dies sehr deutlich zum Ausdruck bringen.						/ 0
33.	Es ist wichtig, daß ich spontan antworte.						/ 0
34.	Wenn ich bei mehr als 3 Fragen überlegen muß, macht das einen schlechten Eindruck.						/ 0

Bestimmt hast Du Dir schon Gedanken gemacht, wie Du Dich in den - 521 -
künftigen Prüfungen verhalten wirst. Darum geht es bei den fol-
genden Items.

Dazu sind 15 Aussagenalternativen formuliert. Zwischen den Alter-
nativen befindet sich eine Skala. Mache Dein Kreuz in die linke
Hälfte, wenn die linke Aussagenalternative zutrifft. Wähle die
rechte Hälfte, wenn die rechte Alternative zutrifft. Kreuze bitte
die "3" an, wenn die Aussage genau zutrifft. Trifft eine der Al-
ternativen eher zu als die andere, so kreuze bitte die "1" oder
"2" an. Die "unentschiedene" Mittelposition - die "0" - wähle bit-
te nur dann, wenn für Dich eine Entscheidung völlig unmöglich ist.
Wie wirst Du Dich in den künftigen Prüfungen verhalten?

Werde phantasievoll den- ken	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde nüchtern den- ken
Werde besonnen vorgehen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde zügig vorgehen
Werde den Gegenstand ganz- heitlich auffassen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde den Gegenstand zergliedern
Werde die Zusammenhänge auflösen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde Zusammenhänge herstellen
Werde leichtfüßig vor- gehen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde hartnäckig vor- gehen
Werde die Sache in den Griff nehmen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde die Sache auf mich wirken lassen
Werde das Verhältnis der Einzelheiten zu einem übergeordneten Ganzen wichtig nehmen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde das Verhältnis der Einzelheiten un- tereinander betonen
Werde rasch vorgehen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde gründlich vor- gehen
Werde systematisch vor- gehen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde ohne strenge Sys- tematik vorgehen
Werde variabel vorgehen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde stetig vorgehen
Werde mich anstrengen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde gelassen bleiben
Werde den Kern der Sache beachten	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde die vielfältigen Einzelheiten der Sache beachten
Werde unbekümmert vor- gehen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde sorgsam vorgehen
Werde Dinge miteinander verknüpfen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde Dinge voneinan- der trennen
Werde Einfälle auf- steigen lassen	3 2 1 0 1 2 3 0--0--0--0--0--0--0	Werde Einfälle unter- drücken

7.5 Listen, Schreiben, Beisitzerblatt

	Seite
1 Seite Liste, in die sich die Kandidaten bei der Verteilung der Prüfungstermine eintragen konnten	523
1 Seite Schreiben an diejenigen Kandidaten, die sich <u>nicht</u> bei der Verteilung der Termine in die Listen eingetragen haben	524
1 Seite Mahnschreiben (ca. 1 Woche nach Versenden des BL-Fragebogens)	525
1 Seite Schreiben an die Beisitzer	526
1 Seite Beisitzerblatt	527

Betr.: Projekt "Prüfungstress" von L.Breitkopf

Name, Vorname (in Druckbuchstaben)	Matrikelnummer	Adresse, unter der Sie in der Prüfungszeit erreichbar sein werden	Vordiplom Diplom (VD/HD)	Wenn Sie prinzipiell bereit sind, an der Untersuchung zuwirken, dann hier unterschreiben

RUHR-UNIVERSITÄT BOCHUM

PSYCHOLOGISCHES INSTITUT

AE Klinische Psychologie

Lothar Breitkopf

463 BOCHUM-QUERENBURG, 8.1.76

UNIVERSITÄTSSTRASSE 150

GEBÄUDE GAFO 04/276

POSTFACH 2148

FERNRUF (02 34) ^{XX} 700/4610

TELEX 08 25 860

Liebe Kandidaten!

Ich beabsichtige, eine empirische Untersuchung für meine Dissertation über "Prüfungsbelastungen" durchzuführen. Ich möchte Euch bitten, daran als Versuchspersonen mitzuwirken. Bis jetzt haben sich alle Vordiplomkandidaten, aber nur 12 Hauptdiplomkandidaten bereit erklärt. Es ist für mich - und wie ich hoffe: für spätere Examinanden - sehr wichtig, daß auch Ihr an der Untersuchung teilnehmt.

Eure Aufgabe wäre es, ca. 14 Tage vor der ersten mündlichen Prüfung einen Fragebogen auszufüllen (Dauer etwa 30 Min.), den Ihr mit der Post erhalten werdet. Ferner wären im Anschluß an 3 Prüfungen einige Ratingskalen auszufüllen (Dauer ca. 10 Min.).

Es ist selbstverständlich, daß alle Daten anonym bleiben, und Rückschlüsse auf Einzelpersonen nicht möglich sein werden. Natürlich wird auch kein Prüfer oder Beisitzer vor, während oder nach der Prüfung die Ergebnisse erfahren. Sämtliche Daten werden von mir bis zum Abschluß der letzten mündlichen Prüfung im Frühjahrs-termin unter Verschuß gehalten.

Für Eure Bemühungen danke ich Euch bereits jetzt. Ich wünsche Euch für die bevorstehenden Prüfungen alles Gute und verbleibe mit freundlichen Grüßen

Lothar Breitkopf

RUHR-UNIVERSITÄT BOCHUM

PSYCHOLOGISCHES INSTITUT

AE Klin. Psychologie
Lothar Breitkopf

463 BOCHUM-QUERENBURG , 22.1.76
UNIVERSITÄTSSTRASSE 150
GEBÄUDE GAFO 04/276
POSTFACH 2148
FERNRUF (0234) 700/4610
TELEX 0825 860

Liebe(r) Prüfungskandidat(in)!

Vor einer Woche hast Du von mir einen Fragebogen im Zusammenhang mit dem Projekt "Prüfungsstress" erhalten. Ich möchte Dich dringend bitten, - wenn noch nicht geschehen - den Fragebogen jetzt auszufüllen und mir noch vor Beginn Deiner mündlichen Prüfungen zurückzusenden.

Wenn Du Gründe hast, die Teilnahme an der Untersuchung zu verweigern, so schicke mir den Fragebogen bitte leer und mit einem Zettel mit Deinem Namen zurück. Du würdest mir damit die organisatorische Arbeit erleichtern, weil ich Dich aus meinem Feldforschungsdesign ausnehmen könnte.

Sollte ich nichts von Dir hören, dann gehe ich davon aus, daß Du an der Untersuchung teilnimmst, den Fragebogen auch ausgefüllt hast, ihn mir aber bewußt erst zu einem späteren Zeitpunkt geben willst.

Ich möchte Dich noch auf die folgende Erleichterung hinweisen: der dem Fragebogen beigelegte Briefumschlag ist frankiert! Du brauchst den Fragebogen nur hineinzustecken und in den nächsten Briefkasten zu werfen.

Solltest Du diesen Brief erhalten, obwohl Du den Fragebogen mir schon hast zukommen lassen, dann betrachte diese Erinnerung (die für mich aber sehr wichtig ist) bitte als gegenstandslos.

Für Deine Bemühungen danke ich Dir!

Lothar Breitkopf

RUHR-UNIVERSITÄT BOCHUM

PSYCHOLOGISCHES INSTITUT
AE Klinische Psychologie
Lothar Breitkopf

463 Bochum, 30.1.76
Ruf~~NR~~ 4610
Postfach 2148

An

.(fast) alle Prüfungsbeisitzer'

Liebe Kollegen!

Im Zusammenhang mit meiner Dissertation möchte ich Sie um einen Gefallen bitten.

Ich führe in der laufenden Prüfungssaison eine Feldstudie über die Belastungen von Prüfungskandidaten durch. Dazu ist es notwendig, bestimmte Variablen des Prüfungsverlaufs zu erfassen.

Ich möchte Sie nun bitten, nach jeder mündlichen Prüfung bei den Prüfern Heckhausen, Hörmann, Schönbach, Schulte und Graefe - sofern Sie Beisitzer sind - den beigelegten Frage-"bogen" auszufüllen. Das Ausfüllen dürfte 30 bis 60 Sekunden dauern.

Herr Graefe als Vorsitzender des Prüfungsausschusses hat dies genehmigt, wenn Sie die Fragen nach Festlegung der Zensur beantworten. Alle Prüfer sind informiert.

Bitte entfernen Sie nach dem Ausfüllen den Zettel mit dem Namen des Kandidaten, sammeln Sie die Frage"bögen" und legen Sie sie dann bitte geschlossen in die Hauspost.

Ich möchte Sie nochmals bitten, mir diesen kleinen Gefallen zu tun. Für Ihre Bemühungen danke ich Ihnen im Voraus.

Mit freundlichen Grüßen

Lothar Breitkopf

Projekt "Prüfungsstress"

L. Breitkopf

Bitte beantworten Sie folgende Fragen zur soeben beendeten Prüfung:

1. Ist der Kandidat beim Antworten ausgewichen durch:

- | | | | | | | | | |
|---|-----------|---|---|---|---|---|---|-------------|
| a) Verwendung von Themen, die inhaltlich zur Prüfung gehören, jedoch nicht gefragt wurden | gar nicht | 0 | 1 | 2 | 3 | 4 | 5 | sehr häufig |
| b) Reden über Gefühle | | 0 | 1 | 2 | 3 | 4 | 5 | |

2. Wie häufig hat der Kandidat Fragen explizit zurückgewiesen oder spontan mit "weiß ich nicht" beantwortet?

- 0 keinmal
0 1 mal
0 2 mal
0 3 mal
0 öfter

3. Hat der Prüfer den Kandidaten immer ausreden lassen?

- 0 ja 0 nein

Wenn nicht, wie häufig hat der Prüfer den Kandidaten unterbrochen?

- 0 1 mal
0 2 mal
0 3-4 mal
0 5-6 mal
0 öfter

4. Konnten Sie beim Prüfer während der Prüfung Gefühle beobachten?
(Zutreffendes bitte ankreuzen!)

- | | | |
|--------|------|----------------------------|
| 0 nein | ja , | 0 Ärger |
| | | 0 Freude |
| | | 0 Verblüffung/Verwunderung |
| | | 0 Angst/Sorge |
| | | 0 Langeweile |

5. Weicht Ihre Einschätzung der Prüfungsleistung des Kandidaten von der endgültigen Zensur ab? (= Ihre Einschätzung vor der Diskussion mit dem Prüfer)

- 0 nein 0 ja

6. Bitte schätzen Sie jetzt noch ab, wie aufgeregt der Kandidat war:

a) zu Beginn,	extrem ruhig	0	1	2	3	4	5	6	7	8	9	extrem aufgeregt
b) am Ende der Prüfung.		0	1	2	3	4	5	6	7	8	9	

7.6 Auswertungsschlüssel und Programme

Die allen Skalen für die Ablochung zugeordneten Werte beginnen links mit '1' und steigen ganzzahlig nach rechts; die Wertzuweisung bei der HiS betrug z. B. 1-2-3-4-5.

HIS

```
001300 C      AUTOMATISCHE AUSWERTUNG DES HIS.
001301 C      NACH EINER TRANSFORMATION WERDEN 2 KENNWERTE GEBILDET:
001302 C      K1=SUMME ALLER ANTWORTEN, K2=SUMME ALLER ANTW. GS. '3'
001303 C      DIE VARIABLEN 7 UND 21 WERDEN IGNORIERT.
001304 C      L.BREITKOPF, 29.6.76
001305 C      INTEGER X(20),Y(20),J,K1,K2
001306 1 READ (1,10,END=99)J,(X(I),I=1,20)
001307 10 FORMAT (15,611,X,1311,X,11)
001308      DO21=1,20
001309      X(1)=(X(1)-6)*X(-1)
001310      Y(1)=0
001311      IF (X(1),GT.2)Y(1)=X(1)-2
001312 2 CONTINUE
001313      K1=0
001314      K2=0
001315      DO31=1,20
001316      K1=K1+X(1)
001317      K2=K2+Y(1)
001318 3 CONTINUE
001319      WRITE (2,20)J,K1,K2,(X(I),I=1,20)
001320 20 FORMAT(15,213,2011,' HIS BREI')
001321      GOTO1
001322 99 STOP 'HIS IST AUSGEWERTET, C*'
001323      END
```

BDI

```

000800 C   AUTOMATISCHE AUSWERTUNG DES BDI1
000801 C   DIE VARIABLE 8 BLEIBT UNBERUECKSICHTIGT.
000802 C   DIE VARIABLEN 3 UND 7 WERDEN NEGATIV GEWICHTET.
000803 C   ES WERDEN 2 KENNWERTE BERECHNET:KENNWERT1 IST SCORESUMME(B1)
000804 C   KENNWERT2 IST X(1) GROESSER 3.
000805 C   L. BREITKOPF, 29.5.76
000806 C   INTEGER X(20),Y(20),B1,B2,J
000807 1 READ (1,10,END=99)J,(X(I),I=1,20)
000808 10 FORMAT (15,711,X,1311)
000809 B1=0
000810 B2=0
000811 DO2 I=1,20
000812 Y(I)=0
000813 C   TRANSFORMATION
000814 X(1)=(X(I)-6)*X(-1)
000815 2 CONTINUE
000816 X(3)=(X(3)-6)*X(-1)
000817 X(7)=(X(7)-6)*X(-1)
000818 DO3 I=1,20
000819 B1=B1+X(I)
000820 Y(I)=X(I)-2
000821 IF (Y(I).LT.0)Y(I)=0
000822 B2=B2+Y(I)
000823 3 CONTINUE
000824 WRITE (2,20)J,B1,B2,(X(I),I=1,20)
000825 20 FORMAT(15,213,2011,4X,' BDI BREI')
000826 GOTO1
000827 99 STOP 'BDI IST BERECHNET, DAS HAST DU FEIN GEMACHT'
000828 END

```

STAI

```

001000 C   AUTOMATISCHE AUSWERTUNG DES STAI (UEBERSETZUNG LAZARUS)
001001 C   L. BREITKOPF, 29.5.76
001002 C   INTEGER J,A,NA,X(20),S
001003 1 READ (1,10,END=99)J,(X(I),I=1,20)
001004 A=0
001005 NA=0
001006 S=0
001007 A=X(3)+X(4)+X(6)+X(7)+X(9)+X(12)+X(13)+X(14)+X(17)+X(18)
001008 NA=X(1)+X(2)+X(5)+X(8)+X(10)+X(11)+X(15)+X(16)+X(19)+X(20)
001009 S=A-NA+50
001010 WRITE (2,20)J,S,A,NA,(X(I),I=1,20)
001011 GOTO 1
001012 10 FORMAT(15,22X,2011)
001013 20 FORMAT(15,312,2011,' STAI: GESAMTSCORE,A,NA,ROHDAT.. BREI')
001014 99 STOP 'STAI IST BERECHNET, C* FUER DICH'
001015 END

```

TVD

```

000500 C   PROGRAMM ZUM SORTIEREN DER TVD- DATEN
000501 C   DAS PROGRAMM MUSS DER AUTOMATISCHEN AUSWERTUNG VORANGEHEN
000502 C   ES WIRD AUSGEGEBEN (DATEI=2): WERDE -A, -V, -U, HABE/BIN
000503 C   -A, -V, -U
000504 C   INTEGER W(15),H(15),J
000505 C   1 READ (1,10,ERR=98,END=99)J,(W(I),I=1,15),(H(I),I=1,15)
000506 10  FORMAT (15,3011)
000507 C   DO2 I=1,15
000508 C   W(I)=W(I)-4
000509 C   2 H(I)=H(I)-4
000510 C   WRITE (2,20)J,W(2),W(5),W(8),W(11),W(13),
000511 C   AW(1),W(6),W(9),W(10),W(15),
000512 C   AW(3),W(4),W(7),W(12),W(14),
000513 C   AH(2),H(5),H(8),H(11),H(13),
000514 C   AH(1),H(6),H(9),H(10),H(15),
000515 C   AH(3),H(4),H(7),H(12),H(14)
000516 20  FORMAT (15,3012,'TVD-SORT',2X,'BREI')
000517 C   GOTO 1
000518 98  STOP 'FEHLER IST AUFGETRETEN'
000519 99  STOP 'ENDE TVD- SORTIERPROGRAMM'
000520 C   END

```

```

000700 C   PROGRAMM ZUR AUTOMATISCHEN AUSWERTUNG DES TVD.
000701 C   VORHER MUSS SORTIERT UND TRANSFORMIERT WERDEN.
000702 C   ES WERDEN NUR DIE MITTELWERTE AUSGEGEBEN.
000703 C   L. BREITKOPF, 31.5.76
000704 C   INTEGER J
000705 C   REAL VH,AH,UH,VW,UW,AW,W(15),H(15)
000706 C   1 READ (1,10,END=99)J,(W(I),I=1,15),(H(I),I=1,15)
000707 10  FORMAT (15,30F2,0)
000708 C   VH=0
000709 C   UH=0
000710 C   AH=0
000711 C   VW=0
000712 C   UW=0
000713 C   AW=0
000714 C   VH=(-H(6)+H(7)+H(8)-H(9)-H(10))/5
000715 C   AH=(-H(1)+H(2)+H(3)-H(4)+H(5))/5
000716 C   UH=(-H(11)+H(12)-H(13)-H(14)-H(15))/5
000717 C   VW=(-W(6)+W(7)+W(8)-W(9)-W(10))/5
000718 C   UW=(-W(1)+W(2)+W(3)-W(4)+W(5))/5
000719 C   UW=(-W(11)+W(12)-W(13)-W(14)-W(15))/5
000720 C   WRITE (2,20)J,AH,VW,UW,AH,VH,UH
000721 20  FORMAT (15,6F5,2,' TVD - MITTELWERTE',20X,'BREI')
000722 C   GOTO 1
000723 99  STOP 'MITTELWERTE SIND BERECHNET'
000724 C   END

```


FBL

```

000010
000020 C AUTOMATISCHE AUSWERTUNG DER FBL(VERSION BREITKOPF)
000030 C GEGENUEBER DER ORIGINALVERSION FEHLEN DIE ITEMS:
000040 C 51,47,44,22,20,11,101111
000050 C DIESES PROGRAMM MUESSTE ALSO BEI VERRECHNUNG DER ORIGINAL-
000060 C VERSION ENTSPRECHEND GEAEENDERT WERDEN.
000070 C L. BREITKOPF, 9.6.76
000080 C INTEGER J
000090 REAL X(71),ALLG,EMOR,HERZ,DARM,KOPF,ANSP,SENS,SCHM,MOTO,HAUT,SUMM
000100 1 READ(1,10,END=99)J,(X(I),I=1,71)
000110 10 FORMAT (15,71F1.0)
000120 DO21=1,71
000130 2 X(I)=X(I)-6)*A(-1)
000140 ALLG=X(1)
000150 - EMOR=0.
000160 HERZ=0.
000170 DARM=0.
000180 KOPF=0.
000190 ANSP=0.
000200 SENS=0.
000210 SCHM=0.
000220 MOTO=0.
000230 HAUT=0.
000240 SUMM=0.
000250 DO301=4,9
000260 30 ALLG=ALLG+X(1)
000270 DO401=10,16
000280 40 EMOR=EMOR+X(1)
000290 DO501=17,22
000300 50 HERZ=HERZ+X(1)
000310 DO601=23,30
000320 60 DARM=DARM+X(1)
000330 DO701=31,38
000340 70 KOPF=KOPF+X(1)
000350 DO801=41,46
000360 80 SENS=SENS+X(1)
000370 DO901=47,54
000380 90 SCHM=SCHM+X(1)
000390 DO1001=58,65
000400 100 MOTO=MOTO+X(1)
000410 DO1101=66,71
000420 HAUT=HAUT+X(1)
000430 110 CONTINUE
000440 DO1201=1,71
000450 120 SUMM=SUMM+X(1)
000460 ANSP=X(2)+X(3)+X(39)+X(40)+X(55)+X(56)+X(57)
000470 ALLG=(ALLG/7)*8
000480 EMOR=(EMOR/7)*8
000490 HERZ=(HERZ/6)*8
000500 ANSP=(ANSP/7)*8
000510 SENS=(SENS/6)*8
000520 SUMM=(SUMM/71)*78
000530 WRITE(2,20)J,ALLG,EMOR,HERZ,DARM,KOPF,ANSP,
000540 *SENS,SCHM,MOTO,HAUT,SUMM
000550 20 FORMAT(15,10F5.2,7F6.2,' FBL BREI')
000560 GOTO 1
000570 99 STOP 'ENDE DER BESCHWERDEN'
000580 END

```

Auswertung der Attribuierungsskalen
der BL-Datenerhebung

```
INTEGER I(12), IV, IK, EV, EK, C, J(12), IV2, IK2, EV2, EK2
IV=0
EV=0
IK=0
EK=0
IV2=0, EV2=0, IK2=0, EK2=0
100READ(1,10,ERR=98,END=99)C,(I(K),K=1,12),(J(K),K=1,12)
10FORMAT(15,20X,2411)
DO11 K=1,12
IF(I(K).LT.3.AND.I(K).NE.0)I(K)=I(K)-3
IF(I(K).GT.2)I(K)=I(K)-2
IF(J(K).LT.3.AND.J(K).NE.0)J(K)=J(K)-3
IF(J(K).GT.2)J(K)=J(K)-2
EV=I(2)-I(6)+I(10)
IV=I(1)-I(5)+I(9)
IK=I(3)-I(7)+I(11)
EK=I(4)-I(8)+I(12)
EV2=J(2)-J(6)+J(10)
IV2=J(1)-J(5)+J(9)
IK2=J(3)-J(7)+J(11)
11EK2=J(4)-J(8)+J(12)
WRITE(2,30)C,IK,EK,IV,EV,IK2,EK2,IV2,EV2,(I(K),K=1,12),(J(K),K=1,12)
2)
30FORMAT(15,812,2412,'ATTR. BRE1')
GOTO100
98STOP 'FEHLER'
99STOP 'ENDE ATTRIBUIERUNG'
END
```

7.7 Liste der in den Konfigurationsfrequenz-
analysen verrechneten Variablen, Angaben zur
Klassenbildung und zur Interpretation der
Klassen

Datensatz	Variable in der KFA	untere Klassen- grenze	Median	obere Klassen- grenze	Interpretation der Stufe	
					1	2
Test- verhaltens- differential (TVD)	Anspannung (Werde)	-3.0	0.9	3.0	wird gelassen vorgehen	wird angespannt vorgehen
	Variabilität (Werde)	-3.0	-0.3	3.0	wird stetig vorgehen	wird variabel vorgehen
	Unifizierung (Werde)	-3.0	1.3	3.0	wird analytisch vorgehen	wird synthetisch vorgehen
	Anspannung (Habe)	-3.0	0.4	3.0	ist gelassen vorgegangen	ist angespannt vorgegangen
	Variabilität (Habe)	-3.0	0.0	3.0	ist stetig vorgegangen	ist variabel vorgegangen
	Unifizierung (Habe)	-3.0	0.9	3.0	ist analytisch vorgegangen	ist synthetisch vorgegangen
Taktiken	Fragen explizit zu- rückgewiesen oder mit "weiß ich nicht" spontan beantwortet	0.0	0.5	6.0	nie	einmal oder mehr- fach
	Prüfer hat unter- brochen	0.0	0.5	7.0	nie	einmal oder mehr- fach
	Aufregung zu Beginn der Prüfung	0.0	4.5	9.0	unterdurch- schnittlich	überdurchschnitt- lich
	Zensureinschätzung des Besitzers	0.0	0.5	1.5	weicht nicht von der des Prüfers ab	weicht von der des Prüfers ab
Emotions- tests	rie	10.0	49.7	99.0	unterdurch- schnittlich	überdurchschnitt- lich
	STAI	10.0	58.7	99.0	unterdurch- schnittlich	überdurchschnitt- lich
	BDI	10.0	51.7	99.0	unterdurch- schnittlich	überdurchschnitt- lich
Analog- skalen	Nachprüfung	0.5	3.4	9.0	zustimmen	ablehnen
	Erschöpfung	0.5	5.1	9.0	voll fit	total erschöpft
	Konzentrations- fähigkeit	0.5	5.3	9.0	hoch	niedrig
Faktor- scores der Beschwerde- liste	BL1 und BL2	5.0	50.0	100.0	unterdurch- schnittlich	überdurchschnitt- lich
Attribuierungs- skalen	IK	-9.0	3.5	9.0	Unfähigkeit	Fähigkeit
	Ex	-9.0	2.8	9.0	schlechter Prüfer	guter Prüfer
	II	-9.0	5.1	9.0	Faulheit	Anstrengung
	EJ	-9.0	1.9	9.0	Pech	Glück
Leistungs- daten	erwartete Zensur	0.5	2.1	5.0	1 oder 2	3 und schlechter
	erhaltene Zensur	0.5	1.9	5.1	1	2 und schlechter
	Prozentsatz des Stoffes	0.0	51.2	74.9	weniger als die Hälfte	mehr als die Hälfte

Datensatz	Variable in der KFA	untere Klassengrenze	Median	obere Klassengrenze	Interpretation der Stufe	
					1	2
Strategien- sammlung	Item 1 (Wenn ich anderer Meinung bin, dann werde ich auch in einer Prüfung einem Professor widersprechen)	0.5	1.5	6.0	"richtig"	"weiß ich nicht" bis "falsch"
	Item 4 (Ich werde möglichst wenig sagen)	0.5	4.5	6.0	"richtig" bis "eher falsch"	"falsch"
	Item 5 (Ich werde mich mutig zeigen)	0.5	2.5	6.0	"richtig"	"weiß ich nicht" bis "falsch"
	Item 8 (Am Anfang bin ich sicherlich aufgeregt, aber das wird sich dann legen)	0.5	1.5	6.0	"richtig"	"eher richtig" bis "falsch"
	Item 14 (Das Beste ist, wenn ich kurz und bündig auf die Fragen des Prüfers antworte)	0.5	2.5	6.0	"richtig"	"weiß ich nicht" bis "falsch"
	Item 17 (Wenn ich auf eine Frage die Antwort nicht weiß, dann heit das, da ich bei der Vorbereitung nicht aufgepat habe)	0.5	1.5	6.0	"richtig" bis "weiß ich nicht"	"falsch"
	Item 22 (Ich gehe davon aus, da mir in der Prüfung auch auf schwere Fragen eine Antwort einfallen wird)	0.5	2.5	6.0	"richtig"	"weiß ich nicht" bis "falsch"
	Item 24 (Prüfer stellen nur scheinbar leichte Fragen, in Wirklichkeit wollen sie mehr wissen)	0.5	2.5	6.0	"richtig"	"weiß ich nicht" bis "falsch"
	Item 26 (Es kommt darauf an, Zusammenhänge zu sehen und zu beschreiben)	0.5	1.5	6.0	"richtig"	"eher richtig" bis "falsch"
	Item 28 (Es kommt darauf an, da ich die Zusammenhänge und Unterschiede anschaulich darstelle)	0.5	1.5	5.0	"richtig"	"eher richtig" bis "falsch"
	Item 31 (Es ist besser, wenn ich nicht allzu kritisch auftrete)	0.5	1.5	6.0	"richtig" bis "weiß ich nicht"	"falsch"
	Item 34 (Wenn ich bei mehr als 3 Fragen überlegen mu, macht das einen schlechten Eindruck)	0.5	1.5	6.0	"richtig" bis "weiß ich nicht"	"falsch"

Folgende Variablen wurden trichotomisiert (1., 25% / 50% / 75%):

- Einschätzung des Prüfungsverlaufs (Analogskala aus AN) mit den Klassengrenzen 0.5, 4., 6.: 1=un-
- (schlechter, 2=genauso wie, 3=besser als erwartet);

- die Differenzen innerhalb des "VO (werde minus habe)" mit den Klassengrenzen für
Anspannung: -3.0, -0.1, 0.6 und 4.0 (1=im Zukunft gelassener, 2=unverändert, 3=im Zukunft angespannter),
Verbilligung: -1.2, -0.5, 0.1 und 4.0 (1=im Zuk. stetiger, 2=unverändert, 3=im Zuk. variabler),
Umfröierung: -5.0, 0.1, 2.7 und 4.0 (1=im Zuk. analytischer, 2=unverändert, 3=im Zuk. stetiger).

7.8 Mittelwerte, Standardabweichungen und
Missing Values der Items der Strategien-
sammlung getrennt nach Datenerhebungs-
zeitpunkten

f 1 - Mittelwerte (M) und Standardabweichungen (s)

Item	Zeitpunkt	Daten- hebungs- Mittelwert			Item	Zeitpunkt	Daten- hebungs- Mittelwert			Item	Zeitpunkt	Daten- hebungs- Mittelwert		
		M	s	MV			M	s	MV			M	s	MV
1	BL	2,63	1,13	1	15	BL	1,93	0,72	-	22	BL	2,56	0,79	2
	P1	2,16	0,67	1		P1	2,00	0,95	-		P1	2,59	0,99	-
	P2	2,16	0,84	-		P2	2,00	1,04	-		P2	2,63	0,96	-
	P3	2,11	0,91	2		P3	1,92	0,88	1		P3	2,54	1,16	2
	P4	2,05	0,95	-		P4	1,84	0,83	1		P4	2,65	1,04	-
2	BL	3,16	1,67	2	16	BL	3,33	1,19	-	23	BL	2,56	1,03	-
	P1	2,81	1,05	-		P1	1,97	0,89	1		P1	2,29	0,84	-
	P2	2,62	1,04	-		P2	2,16	1,11	-		P2	2,10	0,81	-
	P3	2,28	0,91	3		P3	1,97	0,99	2		P3	2,67	1,22	1
	P4	2,74	1,05	1		P4	4,15	1,18	-		P4	2,69	1,22	1
3	BL	2,59	0,82	7	17	BL	3,17	1,20	7	24	BL	3,24	0,97	8
	P1	2,54	0,93	-		P1	3,26	1,14	1		P1	3,24	1,01	5
	P2	2,50	0,95	-		P2	3,16	1,14	1		P2	3,08	1,04	1
	P3	2,25	0,91	3		P3	3,14	1,18	2		P3	3,24	1,05	5
	P4	2,26	0,73	1		P4	2,96	1,36	-		P4	3,53	0,77	1
4	BL	4,08	0,73	1	18	BL	4,37	0,69	14	25	BL	2,78	1,00	1
	P1	4,20	0,81	-		P1	2,27	0,92	-		P1	2,73	1,05	-
	P2	4,05	0,71	-		P2	2,40	0,98	-		P2	2,60	1,08	-
	P3	4,11	0,82	3		P3	2,49	1,04	2		P3	2,34	0,97	1
	P4	4,16	0,87	-		P4	2,30	0,92	-		P4	2,60	0,94	-
5	BL	2,76	1,02	11	19	BL	2,68	1,07	1	26	BL	1,76	0,62	-
	P1	2,64	0,86	5		P1	2,98	1,11	-		P1	1,88	0,75	-
	P2	2,50	0,95	2		P2	2,78	1,03	-		P2	1,85	0,66	-
	P3	2,23	1,00	4		P3	2,71	1,06	1		P3	1,87	0,54	2
	P4	2,58	0,96	1		P4	3,00	1,17	-		P4	1,95	0,61	-
6	BL	2,45	0,86	3	13	BL	2,26	0,71	2	27	BL	2,48	0,91	1
	P1	2,53	0,86	1		P1	2,34	1,01	-		P1	2,59	1,05	-
	P2	2,51	0,89	2		P2	2,43	0,98	-		P2	2,41	0,91	1
	P3	2,55	1,07	5		P3	2,14	0,96	3		P3	2,50	1,06	3
	P4	2,66	0,94	-		P4	2,35	0,88	-		P4	2,26	1,01	-
7	BL	3,29	1,22	12	14	BL	2,85	1,00	1	28	BL	1,68	0,52	-
	P1	3,13	1,19	3		P1	2,98	1,01	-		P1	1,83	0,70	-
	P2	2,87	1,10	2		P2	2,64	1,13	2		P2	1,89	0,47	1
	P3	2,98	1,02	2		P3	2,68	1,13	2		P3	1,84	0,64	1
	P4	2,76	1,17	-		P4	2,85	1,09	-		P4	2,09	0,66	-

LITERATURHINWEISE

- ABRAMOWITZ, S. I.: Locus of control and self-reported depression among college students.
Psychol. Rep., 1969, 25, 149-150.
- ADLER, A.: Menschenkenntnis.
Leipzig, 1958.
- AGER, M.: University examinations and the employment market.
Universities Quarterly, 1966/67, 21, 286-291.
- ALPERT, R. & R. N. HABER: Anxiety in academic achievement situations.
J. abn. soc. Psychol., 1960, 61, 207-215.
- ANISMAN, H. & T. G. WALLER: Effects of inescapable shock on subsequent avoidance performance: role of response repertoire changes.
Behav. Biol., 1973, 9, 331-355.
- ARIETI, S. A.: A re-examination of the phobic symptom and of symbolism in psychopathology.
Amer. J. Psychiat., 1961, 118, 106-110.
- AUSUBEL, D. P.: Psychologie des Unterrichts.
Weinheim, 1974.
- BAINBRIDGE, P. L.: Learning in the rat: Effect of early experience with an unsolvable problem.
J. Comp. Physiol. Psychol., 1973, 82, 301-307.
- BALCH, J.: The influence of the evaluating instruments on student's learning.
Amer. Educ. Res. J., 1964, 1, 169-182.
- BATESON, G., D. D. JACKSON, J. HALEY & J. W. WEAKLAND: Auf dem Wege zu einer Schizophrenie-theorie.
In: BATESON, G., D. D. JACKSON, R. D. LAING u. a. (Hrsg.): Schizophrenie und Familie (Toward a Theory of Schizophrenia). Frankfurt, 1970
Auch in: Behavioral Science, 1956, 1, 251-264.

- BAUMANN, U.: Psychologische Taxometrie: Eine Methodenstudie über Ähnlichkeitskoeffizienten, Q-Clusteranalyse und Q-Faktorenanalyse.
Bern, 1971.
- BECK, A. T.: Depression: Clinical, Experimental and Theoretical Aspects.
New York, 1967.
- BECK, A. T.: Psychoanalytic Research in Depression.
In: FIEVE, R. R. (Ed.): Depression in the 70's - Modern Theory and Research.
Excerpta Medica, 1971.
- BECK, A. T. & A. BEAMESDERFER: Assessment of depression: the depression inventory.
Pharmacopsychiat., 1974, 7, 151-169.
- BECK, A. T., B. SETHI & R. TUTHILL: Childhood bereavement and adult depression.
Arch. Gen. Psychiat., 1963, 9, 295-302.
- BECK, A. T., C. J. WARD, M. MENDELSON, J. MOGK & J. ERBAUGH: An inventory for measuring depression.
Arch. Gen. Psychiat., 1961, 4, 561-571.
- BECK, A. T., A. WEISSMAN, D. LESTER & L. TREXLER: The measurement of pessimism: the hopelessness scale.
J. cons. clin. Psychol., 1974, 42, 861-865.
- BECKMANN, D. & H. E. RICHTER: Gießen-Test (GT).
Bern, 1972.
- BEHREND, E. R. & M. E. BITTERMAN: Sidman avoidance in the fish.
J. exp. Anal. Behavior, 1963, 13, 229-242.
- BELLAK, H. J. & L. SMALL: Kurzpsychotherapie und Notfallpsychotherapie.
Frankfurt, 1972.
- BERINGER, K.: Über hysterische Peaktionen bei Fliegerangriffen.
In: ZWINGMANN, Ch. (Hrsg.): Katastrophenreaktionen. Frankfurt/M., 1971, 18-25.

- BETZ, D.: Phrythmische Schwankungen als Fehler in der Notengebung bei mündlichen Prüfungen.
Psychol. i. Erz. Unterr., 1974, 21, 1-14.
- BEUTEL, P., H. KÜFFNER, E. RÖCK & W. SCHUBÖ:
SPSS - Statistik-Programmsystem für die Sozialwissenschaften - Eine Kurzbeschreibung zur Programmversion 6.
Stuttgart, 1976.
- BINTZ, J.: Between- and within-subject effect of shock intensity on avoidance in goldfish (*Carassius auratus*)
J. comp. physiol. Psychol., 1971, 75, 92-97.
- BIRBAUMER, N.: Unkontrollierbarkeit und Unvorhersagbarkeit.
In: BIRBAUMER, N. (Hrsg.): Psychophysiologie der Angst.
München: 1977, 165-169.
- BIRBAUMER, N., W. TUNNER, R. HÖLZL & L. MITTELSTAEDT:
Ein Gerät zur kontinuierlichen Messung subjektiver Veränderungen.
Z. exp. ang. Psychol., 1973, 20, 173-181.
- BIRCHNELL, J.: The relationship between attempted suicide, depression and parental death.
Brith J. Psychiat., 1970, 116, 307-313.
- BLACK, A. H.: Comments on "Learned Helplessness: Theory and Evidence" by MAIER and SELIGMAN.
J. exp. Psychol: Gen., 1977, 106, 41-43.
- BLINDER, M. G.: The pragmatic classification of depression.
Amer. J. Psychiat., 1966, 123, 259-269.
- BLOOM, B. S., M. D. ENGELHART, E. J. FURST, W. A. HILL & D. R. KRATHWOHL (Eds.):
Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: cognitive domain.
New York, 1956.
Deutsch: Taxonomie von Lernzielen im Kognitiven Bereich. Weinheim, 1974.

- BONIS, M. de: Content analysis of 27 anxiety inventories and rating scales. Mod. Probl. Pharmacopsychiat., 1974, 7, 221-237.
- BORKOVEC, Th. D.: Physiological and Cognitive Processes in the Regulation of Anxiety. In: SCHWARTZ, G. E. & D. SHAPIRO (Eds.): Consciousness and Self-Regulation. Vol. 1. New York, 1976.
- BOYLE, R.: Functional dilemmas in the development of learning. Sociology of Education, 1969, 42, 71.
- BRACEWELL, R. J. & A. H. BLACK: The effects of restraint and noncontingent pre-shock on subsequent escape learning in the rat. Learning and Motivation, 1974, 5, 53-69.
- BRADY, J. V.: Ulcers in "executive" monkeys. Scientific American, 1958, 199, 95-100.
- BRADY, J. V., R. W. PORTER, D. G. CONRAD, J. W. MASON: Avoidance behavior and the development of gastroduodenal ulcers. J. exp. Anal. Behav., 1958, 1, 69-72.
- BRÄUTIGAM, W.: Psychotherapie der Depressionen. In: HIPPIUS, H. & H. SELBACH (Hrsg.): Das depressive Syndrom. München, 1973.
- BRAUDI, W., B. WEPMAN & D. RUSSO: Task and species generality of the "helplessness" phenomenon. Psychonom. Sc., 1969, 16, 154-155.
- BREHM, J. W.: A theory of psychological reactance. New York, 1966.
- BREHM, J. W.: Responses to Loss of Freedom: A Theory of Psychological Reactance. Morriston, N. J., 1972.
- BREITKOPF, L.: Die Hilflosigkeitsskala (HiS). Bochum: Unveröffentl. Manuskript, 1977.
- BREITKOPF, L.: Prognosen mit und Interpretation der Konfigurationsfrequenzanalyse. Bochum: Unveröffentl. Manuskript, 1977(b).

- BRIDGES, P. K.: Practical aspects of the use of some psychological tests of anxiety in a situation of stress.
Brit. J. Psychiat., 1973, 123, 587-596.
- BROADBENT, D. E.: Perception and Communication.
Oxford, 1958.
- BROWN, C. H. & D. Van GELDER: Emotional reactions before examinations.
J. Psychol., 1938, 5, 1-31.
- BURGESS, E. P.: The Modification of Depressive Behaviors.
In: RUBIN, R. D. & C. M. FRANKS (Eds.): Advances in Behavior Therapy.
New York, 1968.
- CAMPBELL, D. T. & D. W. FISKE: Convergent and discriminant validation by the multitrait-multimethod matrix.
Psychol. Bull., 1959, 56, 81-105.
- CAMPBELL, D. T. & J. C. STANLEY: Experimental and quasiexperimental designs for research.
In: GAGE, N. L. (Ed.): Handbook on Research in Teaching.
Chicago, 1963.
Deutsch: INGENKAMP, K. (Hrsg.): Handbuch der Unterrichtsforschung.
Weinheim, 1970.
- CARLIN, A. S. & H. E. ARMSTRONG, Jr.: Aversive conditioning: Learning or dissonance reduction?
J. cons. clin. Psychol., 1968, 32, 674-678.
- CARLTON, P. L. & J. R. VOGEL: Habituation and conditioning.
J. comp. physiol. Psychol., 1967, 63, 348-351.
- CATTELL, R. B. & I. H. SCHEIER: The meaning and measurement of neuroticism and anxiety.
New York, 1961.
- CLAUSS, G. & H. EBNER: Grundlagen der Statistik für Psychologen, Pädagogen und Soziologen.
Berlin, 1968.

- CORTEEN, R. S.: Skin conductance changes and word recall.
Brit. J. Psychol., 1969, 60, 81-84.
- COSTELLO, C. G.: Anxiety and the persisting novelty of input from the autonomic nervous system.
Behav. Ther., 1971, 2, 321-333.
- COSTELLO, C. G.: Depression: loss of reinforcement or loss of reinforcer effectiveness.
Beh. Ther., 1972, 3, 240-247.
- CRONBACH, J. J. & G. C. GLESER: Psychological Tests and Personnel Decisions.
Urbana, 1965.
- DENNEHY, C.: Childhood bereavement and psychiatric illness.
Brit. J. Psychiat., 1966, 112, 1049-1069.
- DOCTOR, R. M. & F. ALTMAN: Worry and emotionality as components of test anxiety: replication and further data.
Psychol. Rep., 1969, 24, 563-568.
- DOHRENWEND, B. P. & B. S. DOHRENWEND: The problem of validity in field studies of psychological disorder.
Intern. J. Psychiat., 1965, 1, 585-610.
- DWECK, C. S.: The role of expectations and attributions in the alleviation of learned helplessness.
J. Pers. Soc. Psychol., 1975, 31, 674-685.
- DWECK, C. S. & N. D. REPUGGI: Learned helplessness and reinforcement responsibility in children.
J. Pers. Soc. Psychol., 1973, 25, 109-116.
- EASTERBROOK, J. A.: The effect of emotion on cue utilization and the organization of behavior.
Psychol. Rev., 1959, 66, 183-201.
- EASTMAN, C.: Behavioral formulations of depression.
Psychol. Rev., 1976, 83, 277-291.

- EKMAN, P. & W. V. FRIESEN: Non-verbal behavior in psychopathology.
In: FRIEDMAN, R. J. & M. M. KATZ (Eds.): The Psychology of Depression - Contemporary Theory and Research. Washington, 1974.
- ELLIS, A.: Reason and Emotion in Psychotherapy. New York, 1962.
- ENGBERT, L. A., G. HANSEN, R. L. WELKER & D. R. THOMAS: Acquisition of key-pecking via auto-shaping as a function of prior experience: "learned laziness"? Science, 1973, 178, 1002-1004.
- ENGEL, G. L.: A life setting conducive to illness: the giving-up-given-up complex. Ann. Internal Med., 1968, 69, 293-300.
- EPSTEIN, S.: The Measurement of Drive and Conflict in Humans: Theory and Experiment. In: JONES, M. R. (Ed.): Nebraska Symposium on Motivation, 1962, 281-321.
- EPSTEIN, S.: Versuch einer Theorie der Angst. In: BIRBAUMER, N. (Hrsg.): Psychophysiologie der Angst. München, 1977, 208-266.
Übersetzt aus: MAHER, B. A.: Progress in Experimental Personality Research. Vol. 4. New York, 1967.
- EPSTEIN, S. & W. D. FENZ: Theory and experiment on the measurement of approach-avoidance conflict. J. Abn. Soc. Psychol., 1962, 64, 97-112.
- EPSTEIN, S. & W. D. FENZ: Steepness of approach and avoidance gradients in humans as a function of experience: theory and experiment. J. exp. Psychol., 1965, 70, 1-12.
- EPTEL, S.: Ein differentialmethodischer Versuch zum Intelligenzproblem. Psychol. Forschung, 1966, 30, 151-210.

- ERTEL, S. & U. SCHINDLER: Intelligenzleistung und Testverhaltensdifferential (a: Teil I, b: Teil II). Diagnostica, 1969, 15, a: 75-98, b: 103-117.
- EYSENCK, H. J. & S. RACHMAN: Neurosen - Ursachen und Heilmethoden, Einführung in die modernen Verhaltenstherapie. Berlin, 1967.
- FAHRENBURG, J.: Die Freiburger Beschwerdenliste FBL. Z. Klin. Psychol., 1976, 1, 79-100.
- FAHRENBURG, J. & P. WALSCHBURGER: Zum Problem der Befindensskalierung in Aktivierungsexperimenten. Hektographiertes Manuskript zur 5. Tagung "Psychophysiologische Methodik". Wien, 1976.
- FENZ, W. D. & S. EPSTEIN: Gradients of physiological arousal of experienced and novice parachutists as a function of an approaching jump. Psychosom. Med., 1967, 29, 33-51.
- FENZ, W. D. & S. EPSTEIN: Specific and general inhibitory reactions associated with mastery of stress. J. exp. Psychol., 1968, 77, 52-66.
- FESTINGER, L.: A theory of social comparison process. Hum. Relat., 1954, 7, 117-140.
- FINN, J. D.: Multivariate - univariate and multivariate analysis of variance, covariance, and regression - A Fortran IV Program, Version V. Buffalo, 1972.
- FISCH, R.: Ein Experiment zur Bestimmung der Stärke von Erfolgshoffung und Prüfungsängstlichkeit bei verschiedenen Zeitdistanzen zum Examen. In: ZIOLKO, H. V. (Hrsg.): Psychische Störungen bei Studenten. Stuttgart, 1969.

- FISCH, R.: Konfliktmotivation und Examen.
Heisenheim, 1970.
- FISCH, R.: Leistungsmotivierung von Prüflingen
bei näherrückendem Examenstermin.
In: ECKSTEIN, B. (Hrsg.): Hochschul-
prüfungen: Rückmeldung oder Repres-
sion?
Blickpunkt Hochschuldidaktik, Hamburg
1971, Bd. 13.
- FLIEGEL, S.: Zur Wirksamkeit von Selbst-Verbali-
sationen bei der Verhaltenstherapie
phobischer Ängste.
Bochum: Unveröffentl. Dissertation,
1978.
- FLUGEL, J. C.: The examination as initiation rite
and anxiety situation.
Intern. J. Psychoanal., 1939, 20,
275-286.
- FORREST, C., R. FRASER & R. PRIEST: Environmental
factors in depressive illness.
Brit. J. Psychiat., 1965, 111, 243-
253.
- FRESE, M. & R. SCHÖFTHALER-RÜHL: Kognitive Ansätze
in der Depressionsforschung.
In: HOFFMANN, N. (Hrsg.): Depressives
Verhalten.
Salzburg, 1976.
- FRITZ, Ch. E. & E. S. MARKS: Die NORC-Untersuchungen
über menschliches Verhalten bei
Katastrophen.
In: ZWINGMANN, Ch. (Hrsg.): Katastro-
phenreaktionen.
Frankfurt/M., 1971, 120-140.
Auch in: J. soc. Issues, 1954, 10,
26-41.
- FRUMKIN, K. & K. H. BROOKSHIRE: Conditioned fear
training and later avoidance learning
in goldfish.
Psychonom. Sc., 1969, 16, 159-160.
- GAENSSLEN, H. & W. SCHUBÖ: Einfache und komplexe sta-
tistische Analyse.
München, 1973.

- GARLINGTON, W. K. & S. B. COTLER: Systematic desensitization of test anxiety.
Beh. Res. Ther., 1968, 6, 247-256.
- GATCHEL, R. J. & J. D. PROCTOR: Physiological correlates of learned helplessness in man.
J. Abn. Psychol., 1976, 85, 27-34.
- GAUDRY, E. & G. D. BRADSHAW: Die unterschiedliche Auswirkung der Angst auf die Leistung bei studienbegleitenden und bei Abschlußprüfungen.
In: KFOHNE, H. W. (Hrsg.): Angst bei Schülern und Studenten.
Hamburg, 1977, 134-137.
Übers. aus: Austral. J. Psychol., 1970, 22, 1-4.
- GEBHARDT, F.: Über die Ähnlichkeit von Faktormatrizen.
Psychol. Beitr., 1968, 10, 591-599.
- GLASS, D. C. & J. E. SINGER: Urban Stress: Experiments on Noise and Social Stressors.
New York, 1972.
- GOLDFRIED, M. R. & T. J. D'ZURILLA: A Behavioral-Analytic Model for Assessing Competence.
In: SPIELBERGER, Ch. D. (Ed.): Current Topics in Clinical and Community Psychology. Vol. 1.
New York, 1969, 151-196.
- GORSUCH, R. L.: Factor Analysis.
Philadelphia, 1974.
- GRAEFE, G. & I. JÖHRENS: Zur Behandlung von Lern- und Leistungsstörungen bei Studierenden.
Studentische Politik, 1971, 3/4,
("Student und Psychotherapie"), 56-59.
- GREEN, W. A., S. GOLDSTEIN & A. J. MOSS: Psychosocial aspects of sudden death: a preliminary report.
Arch. Internal Med., 1972, 129, 725-731.
- GROEGER, W.: Zur Funktion kognitiver Bewältigungsstrategien in Angstsituationen - Eine experimentelle Untersuchung am Beispiel der Angst vor öffentlichem Sprechen. Bochum: Unveröffentl. Dissertation, 1977.

- GROVES, P. M. & R. F. THOMPSON: Habituation: a dual-process-theory.
Psychol. Rev., 1970, 77, 419-450.
- HACKER, W., U. KÖHLER & H. KÜHN: Psychologische Gesichtspunkte zur Eignungsprüfung bei Studienbewerbern.
Wiss. Z. TU Dresden, 1964, 13, 1041-1048.
- HAGGARD, E. A.: Experimental studies in affective processes: I. Some effects of cognitive structure and active participation on certain autonomic reactions during and following experimentally induced stress.
J. exp. Psychol., 1943, 33, 257-284.
- HAKSTIAN, A. R. & S. B. KYUNG: User's Manual to Accompany the Alberta General Factor Analysis Program.
Alberta, 1973.
- HARROW, M. & A. FERRANTE: Locus of control in psychiatric patients.
J. cons. clin. Psychol., 1969, 33, 582-589.
- HARTLAGE, L. C.: Common approaches to the measurement of anxiety.
Amer. J. Psychiat., 1972, 128, 1145-1147.
- HAUSS, K. (Hrsg.): Medizinische Psychologie im Grundriß.
Göttingen, 1976.
- HECKHAUSEN, H.: Leistungsmotivation.
In: THOMAE, H.: Handbuch der Psychologie, 2. Bd. Allgemeine Psychologie, II. Motivation.
Göttingen, 1970.
- HECKHAUSEN, H.: Motive und ihre Entstehung.
In: WEINERT et al. (1974), 133-171.
- HECKHAUSEN, H.: Aufgaben-irrelevante Kognitionen während einer Prüfung: Vorkommen und Störeinfluß.
Bochum: Unveröffentl. Manuskript, 1975.

- HECKHAUSEN, H.: Ein kognitives Motivationsmodell und die Verankerung von Motivkonstrukten. Bochum: Manuskript, 1977.
Erscheint in: LENK, H. (Hrsg.): Handlungstheorien in interdisziplinäre Perspektive (in Vorbereitung).
- HEINERTH, K. H.: Prüfungsangst von Studenten. Psychol. Rundschau, 1972, 23, 79-90.
- HEINERTH, K. H.: Möglichkeiten der Verminderung von Prüfungsangst. Z. Entwicklungspsychol. Päd. Psychol., 1972, 4, 249-260.
- HENN, T. R.: The causes of failure in examinations. Brit. Med. J., 1951, 56, 461-464.
- HILL, O. & J. PRIECE: Childhood bereavement and adult depression. Brit. J. Psychiat., 1967, 113, 743-751.
- HIROTO, D. S.: Locus of control and learned helplessness. J. exp. Psychol., 1974, 102, 187-193.
- HIROTO, D. S. & M. E. P. SELIGMAN: Generality of learned helplessness in man. J. Pers. Soc. Psychol., 1975, 31, 311-327.
- HOFFMANN, N., M. FRESE & J. HARTMANN-ZEILBERGER: Psychologische Therapie bei Depressionen. In: HOFFMANN, N. (Hrsg.): Depressive Verhalten - Psychologische Modelle der Ätiologie und Therapie. Salzburg, 1976.
- HOKANSON, J. E., D. E. De GOOD, S. M. FORREST & T. M. BRITAIN: Availability of avoidance behaviors in modulating vascular-stress responses. J. Pers. Soc. Psychol., 1971, 19, 60-68.
- HORRIDGE, G. A.: Learning of leg position by headless insects. Nature, 1962, 193, 697-698.

- IRLE, M.: Lehrbuch der Sozialpsychologie.
Göttingen, 1975.
- JAMES, W. H. & J. B. ROTTER: Partial and one hundred
per cent reinforcement under chance
and skill conditions.
J. exp. Psychol., 1958, 55, 397-403.
- JANIS, I. L. (a): Emotionale Auswirkungen der Atom-
bombenangriffe auf Hiroshima und Na-
gasaki.
In: ZWINGMANN, Ch. (Hrsg.): Katastro-
phenreaktionen.
Frankfurt/M., 1971, 31-65.
- JANIS, I. L. (b): Luftkrieg und psychische Bela-
stung.
In: ZWINGMANN, Ch. (Hrsg.): Katastro-
phenreaktionen.
Frankfurt/M., 1971, 66-72.
- JOE, V. C.: Review of the internal-external control
construct as a personality variable.
Psychol. Rep., 1971, 28, 619-640.
- KÄPPLER, A.: Einige Zusammenhänge zwischen Studien-
verlauf und psychischer Situation bei
Studenten der Ruhr-Universität Bochum.
Bochum: Unveröffentl. Diplomarbeit,
1977.
- KANFER, F. H. & J. S. PHILLIPS: Lerntheoretische Grund-
lagen der Verhaltenstherapie.
München, 1975.
- KANFER, F. H. & G. SASLOW: Behavioral Analysis.
Arch. Gen. Psychiat., 1965, 529-538.
- KATH, G., Ch. OEHLER & R. REICHWEIN: Studienweg und
Studienerfolg.
Reihe: Studien und Berichte 6.
Berlin, 1966.
- KELLEY, H. H.: Attribution theory in social psychology
In: LEVINE, D. (Ed.): Nebraska Sympo-
sium on Motivation.
Lincoln, 1967.
- KENDELL, R. E.: The continuum model of depressive ill-
ness.
Proc. roy. Soc. Med., 1969, 62, 335-339.

- KEREKJERTO, M. v. & G. A. LIENERT: Depressionsskalen als Forschungsmittel in der Psychopathologie. Pharmakopsychiat. Neuro-Psychopharmakol., 1970, 3, 1-21.
- KIELHOLZ, P.: Diagnose und Therapie der Depressionen für den Praktiker. München, 1971.
- KILOH, L. G. & R. F. GARSIDE: The independence of neurotic and endogenous depression. Intern. J. Psychiat., 1965, 1, 447-465.
- KLEIN, D. C., E. FENCIL-MORSE & M. E. P. SELIGMAN: Learned helplessness, depression and the attribution of failure. J. Pers. Soc. Psychol., 1976, 33, 508-516.
- KLEITER, E.: Voraussetzungskoeffizient und hierarchische Voraussetzungsstrukturanalyse. Zeitschrift für erziehungswissenschaftliche Forschung, 1974, 8, 143-181.
- KLIX, F.: Information und Verhalten - Kybernetische Aspekte der organismischen Informationsverarbeitung - Einführung in naturwissenschaftliche Grundlagen der Allgemeinen Psychologie. Berlin (Ost), 1973.
- KROHNE, H. W.: Schulangst - empirische Befunde, Erklärungsansätze, therapeutische Möglichkeiten. In: KROHNE, H. W. (Hrsg.): Angst bei Schülern und Studenten. Hamburg, 1977.
- KRANTZ, D. S., D. C. GLASS & M. L. SNYDER: Helplessness, stress level, and the coronary prone behavior pattern. J. exp. Soc. Psychol., 1974, 10, 284-300.
- KRAUTH, J. & G. A. LIENERT: KFA - Die Konfigurationsfrequenzanalyse und ihre Anwendung in Psychologie und Medizin. Freiburg/B., 1973.

- KÜHN, H.: Ergebnisse einer Nachuntersuchung über Anfangsschwierigkeiten beim Studium.
Wiss. Z. TU Dresden, 1970, 1, 5-16.
- KÜHN, H. & W. TANNENBERGER: Ergebnisse einer psychologischen Vergleichsuntersuchung zwischen leistungsstarken und leistungsschwachen Studierenden.
Wiss. Z. TU Dresden, 1971, 20, 643-651.
- KVALE, S.: Prüfung und Herrschaft.
Weinheim, 1972.
- LACEY, J. J.: Somatic Response Patterning and Stress: Some Revisions of Activation Theory.
In: APPLE, M. & R. TRUMBULL (Eds): Psychological Stress: Issues in Research.
New York, 1967.
- LACHMAN, S. J.: Psychosomatic Disorders: A Behavioristic Interpretation.
New York, 1972.
- LADER, M. H. & I. M. MARKS: The rating of clinical anxiety.
Acta Psychiat. Scand., 1974, 50, 112-137.
- LAZARUS, A. A.: Learning theory and the treatment of depression.
Behav. Res. & Therapy, 1968, 6, 83-89.
- LAZARUS, R. S.: Psychological stress and the coping process.
New York, 1966.
- LEFCOURT, H. M.: Internal vs. external control of reinforcement: a review.
Psychol. Bull., 1966, 65, 206-220.
- LEGEWIE, H. & L. NUSSELT: Grundlagen und Ergebnisse der Biofeedback-Forschung.
In: LEGEWIE, H. & L. NUSSELT (Hrsg.): Biofeedback-Therapie - Lernmethoden in der Psychosomatik, Neurologie und Rehabilitation.
München, 1975.

- LESSAC, M. & R. L. SOLOMON: Effects of early isolation on the later adaptive behavior of beagles: a methodological demonstration. *Developm. Psychol.*, 1969, 1, 14-25.
- LEVI, L. (Ed.): *Stress and Distress in Response to Psychosocial Stimuli - Laboratory and Real-Life Studies on Sympatho-Adrenomedullary and Related Reactions.* Oxford, 1972.
- LEVIS, D. J.: Learned helplessness: a reply and an alternative S-R interpretation. *J. exp. Psych.: General*, 1976, 105, 47-65.
- LEWINSOHN, P. M.: A Behavioral Approach to Depression. In: FRIEDMAN, R. J. & M. M. KATZ (Eds.): *The Psychology of Depression - Contemporary Theory and Research.* Washington, 1974.
- LEWINSOHN, P. M. & G. E. ATTWOOD: Depression - a clinical research approach. *Psychotherapy*, 1969, 6, 166-171.
- LEWINSOHN, P. M. & J. LIBET: Pleasant events, activity schedules, and depression. *J. abn. Psychol.*, 1972, 79, 291-295.
- LEWINSOHN, P. M. & M. SHAFFER: The use of home observations as an integral part of the treatment of depression - preliminary report and case studies. *J. cons. clin. Psychol.*, 1971, 37, 87-94.
- LEWINSOHN, P. M. & D. A. SHAW: Feedback about interpersonal behavior as an agent of behavior change - a case study in the treatment of depression. *Psychother. Psychosom.*, 1969, 17, 82-88.
- LIBERMAN, R. P. & D. E. RASKIN: Depression: a behavioral formulation. *Arch. Gen. Psychiat.*, 1971, 24, 512-523.
- LIBET, J. M. & P. M. LEWINSOHN: Concept of social skill, with special reference to the behavior of depressed persons. *J. cons. clin. Psychol.*, 1973, 40, 304-312.

- LIEBERT, R. M. & L. W. MORRIS: Cognitive and emotional components of test anxiety: a distinction and some initial data
Psychol. Rep., 1967, 20, 975-978.
- LIENERT, G. A. & M. v. KERKJARTO: Möglichkeiten der Ex-Post-Klassifizierung depressiver Symptome und Patienten mittels Faktoren- und Konfigurationsanalyse.
In: HIPPIUS, H. & SELBACH, S. (Hrsg.): Das depressive Syndrom.
Berlin, 1969, 219-256.
- LINDEN, M.: Depression als aktives Verhalten.
In: HOFMANN, N. (Hrsg.): Depressives Verhalten - Psychologische Modelle der Ätiologie und der Therapie.
Salzburg, 1976, 108-148.
- MAIER, S. F.: Competing motor responses: a reply to BLACK.
J. exp. Psychol.: Gen., 1977, 106, 44-46.
- MAIER, S. F., R. W. ALBIN & T. J. TESTA: Failure to learn to escape in rats previously exposed to inescapable shock depends on nature of escape response.
J. comp. physiol. Psychol., 1973, 85, 581-592.
- MAIER, S. E. & M. E. P. SELIGMAN: Learned helplessness: theory and evidence.
J. exp. Psychol.: General, 1976, 105, 3-46.
- MAIER, S. F., M. E. P. SELIGMAN & R. L. SOLOMON: PAVLOVian fear conditioning and learned helplessness.
In: CAMPBELL, B. A. & R. M. CHURCH (Eds.): Punishment. New York, 1969.
- MAIER, S. F. & T. J. TESTA: Failure to learn to escape by rats previously exposed to inescapable shock is partly produced by associative interference.
J. comp. physiol. Psychol., 1975, 88, 554-564.

- MALLESON, N.: Treatment of pre-examination strain.
Brit. Med. J., 1957, 7, 551-554.
- MALLESON, N.: Avoiding failures by university students.
The Listener, 1959, March, 401-403.
- MANDLER, G.: Helplessness: Theory and Research in Anxiety.
In: SPIELBERGER, Ch. (Ed.): Anxiety-Current Trends in Theory and Research, Vol. II.
New York, 1972.
- MANDLER, G. & D. L. WATSON: Anxiety and the Interruption of Behavior.
In: SPIELBERGER, Ch. (Ed.): Anxiety and Behavior.
New York, 1966, 263-288.
- MARKS, I. M.: Fears and Phobias.
New York, 1969.
- MARKS, I. M. & M. G. GELDER: A controlled retrospective study of behaviour therapy in phobic patients.
Brit. J. Psychiat., 1965, 111, 561-573.
- MASSERMAN, J. H.: Behavior and Neurosis.
Chicago, 1943.
- MAURER-GROELI, Y. A.: Sozialpsychiatrische Untersuchungen bei phobischen und nichtphobischen Neurotikern.
In: BATTEGAY, R. et al. (Hrsg.): Aspekte der Sozialpsychiatrie und Psychohygiene.
Bern, 1975, 115-134.
- MCCULLOCH, T. L. & J. S. BRUNER: The effects of electric shock upon subsequent learning in the rat.
J. Psychol., 1939, 7, 333-336.
- MCDAVID, J. W. & H. HARARI: Social Psychology - Individuals, Groups, Societies.
New York, 1969.

- McLEAN, P. D.: Induced arousal and time of recall as determinants of paired-associate recall.
Brit. J. Psychol., 1969, 60, 57-62.
- McQUITTY, L. L.: Elementary linkage analysis for isolating orthogonal and oblique types and typal relevances.
Educ. ol. Psych. Measmt., 1957, 17, 207-229.
- McQUITTY, L. L.: Capabilities and improvements of linkage analysis as a clustering method.
Educ. ol. Psych. Measmt., 1964, 24, 344-456.
- MENDELS, J.: Depression: The distinction between syndrome and symptom.
Brit. J. Psychiat., 1968, 114, 1549-1554.
- MENDELS, J. & C. COCHRANE: The nosology of depression: the endogeneous-reactive concept.
Amer. J. Psychiat., 1968, 124, 1-11.
- MENDELS, J., N. WEINSTEIN & C. COCHRANE: The relationship between depression and anxiety.
Arch. Gen. Psychiat., 1972, 27, 649-653.
- MENNINGER, W. C.: Psychische Reaktionen auf eine Flutkatastrophe.
In: ZWINGMANN, Ch. (Hrsg.): Katastrophenreaktionen.
Frankfurt, 1971.
- MEPTENS, W.: Sozialpsychologie des Experiments - Das Experiment als soziale Interaktion.
Hamburg, 1975.
- MILLER, G. A., E. GALANTER & K. H. PRIBRAM: Plans and the Structure of Behavior.
New York, 1960.
Deutsch: Pläne und Strukturen des Verhaltens.
Stuttgart, 1973.
- MILLER, W. P.: Psychological deficit in depression.
Psychol. Bull., 1975, 82, 238-260

- MILLER, W. R. & M. E. P. SELIGMAN: Depression and the perception of reinforcement.
J. abn. Psychol., 1973, 82, 62-73.
- MILLER, W. R., M. E. P. SELIGMAN & H. M. KURLANDER: Learned Helplessness, depression, and anxiety.
J. nerv. ment. Dis., 1975, 161, 347-357.
- MOELLER, M. L.: Zur Psychodynamik der neurotischen Prüfungsangst.
In: ZIOLKO, H. U. (Hrsg.): Psychische Störungen bei Studenten.
Stuttgart, 1969.
- MOELLER, M. L. & J. W. SCHEER: Student und psychische Störung - Eine Bestandsaufnahme der psychotherapeutischen Beratung für Studierende an der Justus-Liebig-Universität Gießen.
Gießen: Unveröffentl. Habilitationsschrift, 1970.
- MOELLER, M. L. & J. W. SCHEER: Arbeitsprobleme bei Studenten in der Beratung.
Studentische Politik, 1971, 3/4, "Student und Psychotherapie", 48-55.
- MORRIS, L. W. & R. M. LIEBERT: Effects of anxiety on timed and untimed intelligence tests.
J. cons. clin. Psychol., 1969, 33, 240-244.
- MORRIS, L. W. & R. M. LIEBERT: Relationship of cognitive and emotional components of test anxiety to physiological arousal and academic performance.
J. cons. clin. Psychol., 1970, 35, 332-337.
- MOWRER, O. H.: A stimulus response analysis of anxiety and its role as a reinforcing agent.
Psychol. Rev., 1939, 46, 553-565.
- MOWRER, O. H. & P. VIEK: An experimental analogue of fear from a sense of helplessness.
J. abn. soc. Psychol., 1948, 43, 193-200.

- MÜLLER, H. H. & J. STREKER: Fortran-Programmierungsanleitung.
Mannheim, 1970.
- NAEF, R.: Die Arbeitstechnik von Mittelschülern und Studenten - Ergebnisse einiger Umfragen.
Hektographiertes Manuskript, Institut für Arbeitspsychologie der ETH, Zürich, 1971.
- NEUMANN, J.: Angst und Krankheit vor dem Examen - Wesen, Ursachen, Behebung.
Gütersloh, 1933.
- NIE, N. H., C. H. HULL, J. G. JENKINS, K. STEINBRENNER & D. H. BENT: SPSS - Statistical Package for the Social Sciences - Second Edition.
New York, 1975.
- NISBETT, R. E. & S. SCHACHTER: Cognitive manipulation of pain.
J. exp. soc. Psychol., 1966, 2, 227-236.
- NITSCH, W.: Die soziale Dynamik akademischer Institutionen - Trend-Report zur sozialwissenschaftlichen Hochschulforschung.
Weinheim, 1973.
- OETTING, E. R.: Examination anxiety: Prediction, physiological response and relation to scholastic performance.
J. counsel. Psychol., 1966, 13, 224-228.
- ORNE, M. T.: On the social psychology of the psychological experiment: with particular references to demand characteristics and their implications.
Amer. Psychologist, 1962, 17, 776-783.
- OSGOOD, Ch. E., G. SUCI & P. H. TANNENBAUM: The Measurement of Meaning.
Urbana 1957 and Chicago 1963.
- OVERMIER, J. E.: Interference with avoidance behavior: Failure to avoid traumatic shock.
J. exp. Psychol., 1968, 78, 340-343.

- OVERMIER, J. B. & M. E. P. SELIGMAN: Effects of inescapable shock upon subsequent escape and avoidance learning. J. comp. physiol. Psychol., 1967, 63, 23-33.
- PADILLA, A. M.: Effects of prior and interpolated shock exposures on subsequent avoidance learning by goldfish. Psychol. Rep., 1973, 32, 451-456.
- PADILLA, A. M., C. PADILLA, T. KETTERER & D. GIACOLONE: Inescapable shocks and subsequent avoidance conditioning in goldfish. Psychonom. Sc., 1970, 20, 295-296.
- PAYKEL, E. S., J. K. MYERS, M. V. DIENELT, G. L. KLERNAN, J. J. LINDENTHAL & M. P. PEPPER: Live events and depression. Arch. Gen. Psychiat., 1969, 21, 753-760.
- PAYNE, R. W. & J. H. G. HEWLETT: Thought disorder in psychotic patients. In: EYSENCK, H. J. (Ed.): Experiments in Personality. Vol. 2. London, 1960.
- PERVIN, L. A.: The need to predict and control under conditions of threat. J. Pers., 1963, 31, 570-585.
- PHARES, E. J.: Expectancy change in chance and skill situations. J. abn. Psychol., 1957, 54, 339-342.
- PIEHL, J.: Angst und Konflikt - Experimentelle Untersuchung an Examenskandidaten. Göttingen: Dissertation, 1972.
- PIEHL, J. Bedingungen unterschiedlicher Ursachen-erklärungen von Examensnoten. Z. Entwicklungspsychol. Päd. Psychol., 1976, 8, 54-57.
- PINCKNEY, G.: Avoidance learning in fish as a function of prior fear conditioning. Psychol. Rep., 1967, 20, 71-74.

- PÜLDINGER, W., P. BLASER & A. GEHRING: Zur Quantifizierung psychopathologischer und somatischer Symptome bei depressiven Verstimmungszuständen.
In: HIPPIUS, H. & S. SELBACH (Hrsg.): Das depressive Syndrom.
München, 1969, 291-304.
- POSER, E. G.: Toward a theory of 'behavioral prophylaxis'.
J. Beh. Ther. exp. Psychiat., 1970, 1, 39-43.
- PUCHTA, H. G. & H. U. ZIOLKO: Klinische Erfahrungen mit Prüfungsversagern.
In: ZIOLKO, H. U. (Hrsg.): Psychische Störungen bei Studenten.
Stuttgart, 1969.
- RACHMAN, S. & J. B. BERGOLD: Verhaltenstherapie bei Phobikern.
München, 1976.
- RICHTER, C. F.: Sudden Death Phenomenon in Animals and Humans.
In: FEIFEL, H. (Ed.): The Meaning of Death.
New York, 1959.
- ROEDER, B.: Die Konfigurationsfrequenzanalyse (KFA) nach KRAUTH und LIENERT - Ein handliches Verfahren zur Verarbeitung sozialwissenschaftlicher Daten, demonstriert an einem Beispiel.
Kölner Z. Soziol. Sozialpsychol., 1974, 26, 819-844.
- ROEDER, B.: Parameterfreier Vergleich verschiedener Konfigurationsfrequenzanalysen.
Kölner Z. Soziol. Sozialpsychol., 1976, 2, 152-155.
- ROSS, L. D., J. RODIN & P. G. ZIMBARDO: Toward an attribution therapy: the reduction of fear through induced cognitive-emotional misattribution.
J. Pers. Soc. Psychol., 1969, 12, 279-288.

- ROTH, S. & P. R. BOOTZIN: The effects of experimentally induced expectancies of external control: an investigation of learned helplessness.
J. Pers. Soc. Psychol., 1974, 29, 253-264.
- ROTH, S. & L. KUBAL: Effects of noncontingent reinforcement on tasks of differing importance: facilitation and learned helplessness.
J. Pers. Soc. Psychol., 1975, 32, 680-691.
- POTTER, J. B.: Social Learning and Clinical Psychology.
Englewood Cliffs, N. J., 1954.
- POTTER, J. B.: Generalized expectancies for internal vs. external control of reinforcement
Psychol. Monographs, 1966, 80 (1, Whole No. 609).
- ROTTER, J. B., S. LIVERANT & D. P. CROWNE: The growth and extinction of expectancies in chance controlled and skilled tasks.
J. Psychol., 1961, 52, 161-177.
- ROTTER, J. B. & R. C. MULRY: Internal vs. external control of reinforcement and decision time.
J. Pers. Soc. Psychol., 1965, 2, 598-604.
- SARASON, I. G.: Empirische Befunde und theoretische Probleme bei der Verwendung von Angstskalen.
In: KROHNE, H. W. (Hrsg.): Angst bei Schülern und Studenten.
Hamburg, 1977, 50-60.
- SARGENT, W. & H. J. SHORVON: Acute war neurosis, special reference to PAVLOV's experimental observations and the mechanism of abreaction.
Arch. Neurol. Psychiat., 1945, 54, 231.
- SELIGMAN, M. E. P.: Chronic fear produced by unpredictable electric shocks.
J. comp. physiol., Psychol., 1968, 66, 402-411.

- SELIGMAN, M. E. P.: Fall into helplessness.
Psychology today, 1973, June, 43-48.
- SELIGMAN, M. E. P.: Depression and Learned Helplessness.
In: FRIEDMAN, R. J. & M. M. KATZ (Eds.):
The Psychology of Depression: Contemporary
Theory and Research.
Washington, 1974.
- SELIGMAN, M. E. P. & G. BEAGLEY: Learned helplessness
in the rat.
J. comp. physiol. Psychol., 1975, 88,
534-541.
- SELIGMAN, M. E. P. & D. GROVES: Non-transient learned
helplessness.
Psychonom. Sc., 1970, 19, 191-192.
- SELIGMAN, M. E. P., D. C. KLEIN & W. R. MILLER:
Depression.
In: LEITENBERG, H.: Handbook of
Behavior Therapy,
New York, 1976.
- SELIGMAN, M. E. P. & S. F. MAIER: Failure to escape
traumatic shock.
J. exp. Psychol., 1967, 74, 1-9.
- SELIGMAN, M. E. P., S. F. MAIER & J. GEER: The alleviation
of learned helplessness in the dog.
J. abn. soc. Psychol., 1968, 73,
256-262.
- SELIGMAN, M. E. P., S. F. MAIER & R. L. SOLOMON:
Unpredictable and Uncontrollable
Aversive Events.
In: BRUSH, F. R. (Ed.): Aversive conditioning
and Learning.
New York, 1971, 347-400.
- SELIGMAN, M. E. P. & B. MEYER: Chronic fear and ulcers
as a function of the unpredictability
of safety.
J. comp. physiol. Psychol., 1970, 73,
202-207.
- SELIGMAN, M. E. P., R. A. ROSSELINI & M. KOZAK:
Learned helplessness in the rat:
reversibility, time course, and
immunization.
J. comp. physiol. Psychol., 1975,
88, 542-547.

- SEYLE, H.: The Stress of Life.
New York, 1956.
- SEWARD, J. P. & G. L. HUMPHREY: Avoidance learning
as a function of pretraining in the
cat.
J. comp. physiol. Psychol., 1967,
63, 338-341.
- SHAPIRO, M. B.: The single case in clinical-psychologic-
al research.
J. Gener. Psychol., 1966, 74, 3-23.
- SHERROD, D. R. & R. DOWNS: Environmental determinants
of altruism: the effects of stimulus
overload and perceived control on
helping.
J. exp. soc. Psychol., 1974, 10,
468-479.
- SIMSON, T.: Psychische und psychotische Reaktionen
Erwachsener und Kinder bei Erdbeben.
In: ZWINGMANN, Ch. (Hrsg.): Kata-
strophenreaktionen.
Frankfurt/M., 1971, 99-113.
- SMITH, D. B. D., M. A. WENGER: Changes in automatic
balance during phasic anxiety.
Psychophysiology, 1965, 1, 267-271.
- SMYTHIES, J. R., A. COPPEN & U. KREITMAN: Biologische
Psychiatrie-Entwicklung - Fortschrit-
te - Ausblicke.
Stuttgart, 1970.
- SNAITH, R. P.: A clinical investigation of phobias.
Brit. J. Psychiat., 1968, 114, 673-
697.
- SOKOLOV, E. N.: Higher nervous functions: the orient-
ing reflex.
Annual Rev. Physiol., 1963, 25, 545-
580.
- SOKOLOV, J. N.: Die reflektorischen Grundlagen der
Wahrnehmung.
In: HIEBSCH, H. (Hrsg.): Ergebnisse
der sowjetischen Psychologie.
Berlin, 1967.

- SORGATZ, H.: Motorisches Lernen und Aktivierung.
Bochum: Unveröffentl. Dissertation, 1974.
- SPIELBERGER, Ch. D.: Anxiety and Behavior.
New York, 1966.
- SPIELBERGER, Ch. D., R. L. GORSUCH & R. E. LUSHENE:
STAI Manual for the STATE-TRAIT-
ANXIETY-INVENTORY ("Self-Evaluation-
Questionnaire").
Palo Alto, 1970.
- SUAREZ, S.: Faktorenstruktur der Arbeitstechnik
und Zusammenhänge mit den Schulleistungen
bei 16-19jährigen Mittelschülern.
Hektographiertes Manuskript;
Zürich, Institut für Arbeitspsychologie,
1971.
- SCHACHTER, S. & J. E. SINGER: Cognitive, social and
physiological determinants of emotional
state.
Psychol. Rev., 1962, 69, 379-399.
- SCHALLBERGER, U.: Studienverlauf und Studienerfolg.
Basel, 1974.
- SCHAPIRA, K., T. A. KERR & M. ROTH: Phobias and
affective illness.
Brit. J. Psychiat., 1970, 117, 25-32
- SCHEER, J. W. & H. ZENZ: Studenten in der Prüfung -
Eine Untersuchung zur akademischen
Initiationskultur.
Frankfurt, 1973.
- SCHILDKPAUT, J. J.: The catecholamine hypothesis of
affective disorder; a review of
supporting evidence.
Amer. J. Psychiat., 1965, 112, 509-522.
- SCHLESS, A. P., J. MENDELS, A. KIPPERMAN & C. COCHRANE:
Depression and hostility.
J. nerv. ment. Dis., 1974, 159, 91-100.
- SCHMALE, A. H.: Given up as a final common pathway
to changes in health.
Psychosom. Med., 1971, 8, 18-28.
- SCHNEIDER, K.: Klinische Psychopathologie.
Stuttgart, 1971⁹

- SCHÖNPFLUG, W.: Vorgänge psychophysiologischer Aktivierung und drei Ansätze zu ihrer Untersuchung.
In: SCHÖNPFLUG, W. (Hrsg.): Methoden der Aktivierungsforschung.
Bern, 1969, 11-26.
- SCHULTE, D.: Die Dimensionalität des Feldabhängigkeitskonzepts - ein methodenkritischer Beitrag zur differentiellen Psychologie.
Münster: Unveröffentl. Dissertation, 1972.
- SCHULTE, D.: Grundlagen der Verhaltenstherapie.
In: PONGRATZ, L. J. (Hrsg.): Klinische Psychologie. Handbuch der Psychologie. Bd. 8, 1. Halbb.
Göttingen, 1976.
- SCHULTE, W. & R. TOELLE: Psychiatrie.
Berlin, 1973.
- STENGEL, E.: Prüfungsangst und Prüfungsneurose.
Z. Psychoanal. Pädagog., 1936, 10, 300-320.
- STENGEL, E.: Prüfungsangst und Prüfungsneurose.
Almanach der Psychoanalyse, 1938, 13, 164-193.
- STORMS, M. D. & R. E. NISBETT: Insomnia and the attribution process.
J. Pers. Soc. Psychol., 1970, 16, 319-328.
- TAULBEE, E. S. & H. W. WRIGHT: A psycho-social-behavioral model of therapeutic intervention.
In: SPIELBERGER, C. D. (Ed.): Current Topics in Clinical and Community Psychology III.
New York, 1971.
- TAUSCH, R.: Gesprächspsychotherapie.
Göttingen, 1970.
- TEICHNER, W. H.: Interaction of behavioral and physiological stress reactions.
Psychol. Rev., 1968, 75, 271-291.
- TENNEN, H. & S. J. ELLER: Attributional components of learned helplessness and facilitation.
J. Pers. Soc. Psychol., 1977, 35, 265-271.

- THORNTON, J. W. & P. D. JACOBS: Learned helplessness in human subjects.
J. exp. Psychol., 1971, 87, 367-372.
- THORNTON, J. W. & G. D. POWELL: Immunization to and alleviation of learned helplessness in man.
Amer. J. Psychol., 1974, 87, 351-367.
- THROOP, W. F. & A. P. MacDONALD, Jr.: Internal-external locus of control: a bibliography.
Psychol. Rep., 1971, 28, 175-190.
- TURKSON, S. N. A.: Psychosomatische Erhebungen an Prüflingen.
Gießen: Med. Dissertation (unveröffentl.) 1970.
- VALINS, S. & R. E. NISBETT: Some Implications of Attribution Processes for the Development and Treatment of Emotional Disorders.
New York, 1971.
- WATERS, W. F.: Sympathetic Recovery in Defensive Responding.
J. abn. Psychol., 1970, 76, 355-360.
- WATSON, J. B. & R. RAYNER: Conditioned emotional reactions.
J. exp. Psychol., 1920, 3, 1-14.
- WATZLAWICK, P., J. H. BEAVIN & D. D. JACKSON: Menschliche Kommunikation - Formen, Störungen, Paradoxien.
Bern, 1969.
- WEAKLAND, J. H.: "Double-Bind"-Hypothese and Dreier-Beziehung.
In: BATESON, G., D. D. JACKSON, R. D. LAING u. a.: Schizophrenie und Familie.
Frankfurt, 1970.
- WEINER, B., H. HECKHAUSEN, W. MEYER & R. E. COOK: Causal ascriptions and achievement behavior: the conceptual analysis of effort.
J. Pers. Soc. Psychol., 1972, 21, 239-248.

- WEINER, B., J. FRIEZE, A. KUKLA, L. REED, S. REST & R. M. ROSENBAUM: Perceiving the causes of success and failure. Morristown, N. J., 1971.
- WEINER, B. & A. KUKLA: An attributional analysis of achievement motivation. J. Pers. Soc. Psychol., 1970, 15, 1-20.
- WEINERT, F. E., C. F. GRAUMANN, H. HECKHAUSEN & M. HOFER (Hrsg.): Funk-Kolleg Pädagogische Psychologie, Band I. Frankfurt/M., 1974.
- WEISS, J. M.: Somatic effects of predictable and unpredictable shock. Psychosom. Med., 1970, 32, 397-408.
- WEISS, J. M.: Effects of coping behavior in different warning signal conditions on stress pathology in rats. J. comp. physiol. Psychol., 1971, 77, 1-13.
- WEISS, J. M.: Effects of punishing the coping response (conflict) on stress pathology in rats. J. comp. physiol. Psychol., 1971, 77, 14-21.
- WEISS, J. M.: Effects of coping behavior with and without a feedback signal on stress pathology in rats. J. comp. physiol. Psychol., 1971, 77, 22-30.
- WEISS, J. M.: Psychological factors in stress and disease. Sci. Amer., 1972, 226, 104-113.
- WEISS, J. M., E. E. KPIECKHAUS & R. CONTE: Effects of fear conditioning on subsequent avoidance behavior. J. comp. physiol. Psychol., 1958, 65, 413-421.
- WEISS, J. M., E. A. STONE & N. HARRELL: Coping behavior and brain norepinephrine in rats. J. comp. physiol. Psychol., 1970, 72, 153-160.

- WHITE, R. W.: Motivation reconsidered: the concept of competence.
Psychol. Rev., 1959, 66, 297-333.
- WILLIAMS, C. B. & J. B. NICKELS: Internal-external control-dimension as related to accident and suicide proneness
J. cons. clin. Psychol., 1969, 33, 485-493.
- WINE, J.: Test anxiety and direction of attention.
Psychol. Bull., 1971, 76, 91-104.
- WOLPE, J.: Neurotic depression: experimental analogue, clinical syndromes, and treatment.
Amer. J. Psychotherapy, 1971, 25, 362-368.
- WORTMAN, C. B. & J. W. BREHM: Responses to Uncontrollable Outcomes: An Integration of Reactance Theory and the Learned Helplessness Model.
In: BERKOWITZ, L. E. (Ed.): Advances in Experimental Social Psychology
New York, 1975.
- WORTMAN, C. B., L. PANCIERA, L. SHUSTERMAN & J. HIBSCHUSER: Attributions of causality and reactions to uncontrollable outcomes.
J. exp. soc. Psychol., 1976, 12, 301-316.
- WUNDER, M.: Die Wohnsituation der RUB-Studenten und ihre psychischen Folgen.
Bochum: Unveröffentl. Diplomarbeit, 1977.
- YATES, A. J.: Behavior Therapy.
New York, 1970.
- ZENZ, H. & J. W. SCHEER: Das Examen: Lohn der Angst? Eine Untersuchung über Inhalte und Auswirkung der Prüfungsangst.
Analysen, 1973, 10, 24-27.
- ZIELINSKI, K. & S. SOLTSYK: The effect of pretraining on the acquisition and extinction of the avoidance reflex.
Acta Biologicae Experimentalis, 1964, 24, 73-87.

- ZIMBARDO, P. G.: The Human Choice: Individuation, Reason, and Order versus Deindividuation, Impulse and Chaos.
In: ARNOLD, W. J. & D. LEVINE (Eds.): Nebraska Symposium on Motivation, 1969, 237-307.
- ZIOLKO, H. U.: Melancholie und Erschöpfung bei Studenten.
Psychiat. clin., 1968, 1, 265-268.
- ZIOLKO, H. U.: Echte Probleme der akademischen Jugend.
Ärztliche Praxis, 1970, 84.
- ZUBIN, J. & J. FLEISS: Current Biometric Approaches to Depression.
In: FIEVE, R. R. (Ed.): Depression in the 70's - Modern Theory and Research.
Excerpta Medica, 1971.
- ZUCKERMAN, M., H. PERSKY, K. M. ECKMAN & T. R. HOPKINS: A multitrait multimethod measurement approach to the traits (or states) of anxiety, depression, and hostility.
J. project. Techniq., 1967, 31, 39-48.
- ZUNG, W. W. K.: A self-rating depression scale.
Arch. gen. Psychiat., 1965, 12, 63-70.
- ZUNG, W. W. K.: A rating instrument for anxiety disorders.
Psychosomatics, 1971, 12, 371-379.
- ZUNG, W. W. K.: The differentiation of anxiety and depressive disorders: a biometric approach.
Psychosomatics, 1971, 12, 380-384.
- ZUNG, W. W. K.: The measurement of affects: depression and anxiety.
Mod. Prob. Pharmacopsychiat., 1974, 7, 170-188.
- ZWINGMANN, Ch.: Katastrophenreaktion - Übersicht und sozialkritische Betrachtung.
In: ZWINGMANN, Ch. (Hrsg.): Katastrophenreaktionen.
Frankfurt/M., 1971.

LEBENS LAUF

Name: Lothar Breitkopf

Geburtstag und -ort: 15.03.1949 in Sarstedt, Kreis
Hildesheim-Marienburg

Schulen: Ostern 1955 eingeschult in Hildesheim
Ostern 1959 Wechsel zum mathematisch-
naturwissenschaftlichen, neusprachli-
chen Scharnhorstgymnasium in Hildes-
heim
Sommer 1967 Abitur

Wehrdienst: Oktober 1967 bis März 1969 beim Bun-
desgrenzschutz in Goslar

Studium: SS 1969 Immatrikulation als Student
der Psychologie in Göttingen
SS 1971 Vordiplomprüfung in Göttingen
WS 1971/72 Wechsel an die Ruhr-Univer-
sität Bochum
01.03.73 bis Sommer 1974 studentische
Hilfskraft im Studienbüro der Ruhr-
Universität Bochum
Oktober 1974 Hauptdiplomprüfung in
Bochum

Post-
graduierten-
tätigkeiten: 01.11.74 Wiss. Ass. (mdVb) in der
AE Klinische Psychologie, Prof. Dr.
D. Schulte, Ruhr-Universität Bochum,
ab 01.10.75 Wiss. Hilfskraft und ab
01.04.76 Bezieher eines Stipendiums
nach dem Graduiertenförderungsgesetz;
seit dem 01.09.77 Wiss. Ass. (mdVb)
am Lehrstuhl Medizinische Psychologie,
Prof. Dr. G. Krüskemper, Ruhr-Univer-
sität Bochum

Druck: D. Gräbner, 8602 Altendorf - Tel. 09545 / 8791